

Lösungsmethoden für Variationsungleichungen

D I S S E R T A T I O N

zur Erlangung des akademischen Grades
doctor rerum naturalium
(dr. rer. nat.)
im Fach Mathematik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät II
Humboldt-Universität zu Berlin

von
Dipl.-Math. Andrej Ponomarenko
geboren am 03.10.1974 in Chabarowsk (Russland)

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Jürgen Mlynek

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät II:
Prof. Dr. Elmar Kulke

Gutachter:

1. Prof. Dr. D. Klatte
2. Prof. Dr. C. Großmann
3. Prof. Dr. B. Kummer

eingereicht am: 8. August 2002
Tag der mündlichen Prüfung: 31. Januar 2003

Zusammenfassung

Diese Arbeit ist ein Versuch, verschiedene klassische und neuere Methoden der glatten bzw. nicht-glatten Optimierung zu verallgemeinern und in ihrem Zusammenhang darzustellen. Als Hauptinstrument erweist sich dabei die sogenannte verallgemeinerte Kojima-Funktion. Neben reichlichen Beispielen setzen wir einen besonderen Akzent auf die Betrachtung von Variationsungleichungen, Komplementaritätsaufgaben und der Standardaufgabe der mathematischen Programmierung

$$\min_{x \in \mathbb{R}^n} \{f(x) \mid g_i(x) \leq 0, \quad i = 1, \dots, m\}.$$

Unter natürlichen Voraussetzungen an diese Probleme kann man u.a. Barriere-, Straf- und SQP-Typ-Methoden, die auf Newton-Verfahren basieren, aber auch Modelle, die sogenannte NCP-Funktionen benutzen, mittels spezieller Störungen der Kojima-Funktion exakt modellieren. Daneben werden durch explizite und natürliche Wahl der Störungsparameter auch neue Methoden dieser Arten vorgeschlagen. Die Vorteile solcher Modellierung sind überzeugend vor allem wegen der direkt möglichen (auf Stabilitätseigenschaften der Kojima-Gleichung beruhenden) Lösungsabschätzungen und weil die entsprechenden Nullstellen ziemlich einfach als Lösungen bekannter Ersatzprobleme interpretiert werden können.

Ein weiterer Aspekt der Arbeit besteht in der genaueren Untersuchung der "nichtglatten Fälle". Hier wird die Theorie von verschiedenen verallgemeinerten Ableitungen und dadurch entstehenden verallgemeinerten Newton-Verfahren, die im Buch [14] von B. Kummer und D. Klatte vorgeschlagen und untersucht wurde, intensiv benutzt. Entscheidend ist dabei, dass die benutzten verallgemeinerten Ableitungen auch praktisch angewandt werden können, da man sie exakt ausrechnen kann.

An dieser Stelle möchte ich mich bei meinem wissenschaftlichen Betreuer Prof. Dr. Bernd Kummer herzlich bedanken für die intensive Zusammenarbeit, vielseitige Unterstützung und wertvolle Hilfe, ohne die das Entstehen dieser Arbeit unvorstellbar gewesen wäre.

Außerdem gilt mein besonderer Dank der Unterstützung durch das Graduiertenkolleg "Geometrie und nichtlineare Analysis".

Inhaltsverzeichnis

1	Einführung	7
2	Verallgemeinerte Kojima-Funktion	11
2.1	Kojima-Funktion	11
2.2	Verallgemeinerte Kojima-Funktion	12
2.3	Beispiele	13
2.3.1	Komplementaritätsprobleme	13
2.3.2	Nash-Gleichgewicht im Stackelberg-Spiel	14
2.3.3	Eine verallgemeinerte Gleichung	15
3	Ableitungen und Regularität	17
3.1	Regularität und $TF, CF, \partial F$	17
3.2	Ableitungen für Newton-Verfahren	19
3.2.1	Generalized Jacobian	19
3.2.2	Limes von Jacobians von C^1 -Punkten	20
3.3	Anwendungen auf die Kojima-Funktion	21
3.3.1	Die Ableitungen $\partial F, TF$ und CF	21
3.3.2	Die Ableitungen $\partial_0 F$ und $D^0 F$	23
4	Newton-Verfahren	25
4.1	Allgemeine Konvergenzaussagen	25
4.2	Approximation von $GF(x, u)$ und $F(x)$	27
4.3	Die Anwendung von $D^0 F, LocPC^1$ und $F'(x; u)$	29
4.3.1	$D^0 F$ und $LocPC^1$	29
4.3.2	Verknüpfte Funktionen	31
5	Gezielte Störungen von Kojima-Funktionen	33
5.1	Beziehung zu den bekannten Verfahren	34
5.2	Mix-Methode	37
5.3	Symmetrische Störung	37
5.4	Abschätzungen	39
5.5	Anwendung gestörter Kojima-Funktionen im Newton-Verfahren	40

6	Einheitliche Darstellung von NCP-Ansätzen	43
6.1	NCP-Funktion	43
6.2	Transformation der Iterierten beim Newton-Verfahren	46
6.3	Symmetrische Störung des Kojima-Systems und $pNCP$ -Funktionen im Newton-Verfahren	48
6.3.1	Beliebige $pNCP$ -Funktionen	49
6.3.2	Vereinfachung für streng monotone G aus $pNCP$	53
7	Modelle für Variationsungleichungen	55
7.1	Einheitliche Darstellung des Newton-Verfahrens	56
7.2	Verlauf der Verfahren	60
7.2.1	Gradienten-Verfahren und Monotonie	60
7.2.2	Newton-Verfahren	61

Kapitel 1

Einführung

Der Hauptgegenstand dieser Arbeit sind klassische Variationsungleichungen in endlicher Dimension:

$$\text{Finde ein } x \in C, \text{ so dass } \langle p(x), c - x \rangle \leq 0 \quad \forall c \in C.$$

Dabei ist $C \subset \mathbb{R}^n$ eine nichtleere, abgeschlossene und konvexe Menge und $p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine lokal Lipschitz-stetige Funktion. Die Aufgabe wird oft in der folgenden Form geschrieben:

$$\text{Finde ein } x \in C, \text{ so dass } p(x) \in N_C(x), \quad (1.1)$$

wobei

$$N_C(x) := \{\xi \in \mathbb{R}^n \mid \langle \xi, c - x \rangle \leq 0, \quad \forall c \in C\}.$$

Die Menge $N_C(x)$ heißt Normalenkegel (im Sinne der konvexen Analysis, [27]) von C im Punkt x . Ist $x \notin C$, so sei $N_C(x) := \emptyset$. Wird N_C ersetzt durch eine mehrwertige Abbildung $\Gamma : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, so spricht man von einer verallgemeinerten Gleichung.

Zur numerischen Untersuchung von (1.1) benötigt man eine analytische Beschreibung von C . Sei zunächst

$$C := \{x \mid g_i(x) \leq 0, i = 1, \dots, m\}, \quad g_i \in C^1, \quad g_i \text{ konvex} \quad \forall i. \quad (1.2)$$

Unter einer Regularitäts-Voraussetzung im Punkt x (MFCQ, Abadi-Bedingung oder Calmness, siehe [14]) hat dann der Normalenkegel eine spezielle Form:

$$N_C(x) = \{y^+ Dg(x) \mid g(x) = y^-, y \in \mathbb{R}^m\},$$

wobei $y_i^+ = \max\{0, y_i\}$ und $y_i^- = \min\{0, y_i\}$. Damit wird (1.1) äquivalent zu der Gleichung

$$\begin{aligned} F_1(x, y) &:= -p(x) + y^+ Dg(x) = 0 \\ F_2(x, y) &:= g(x) - y^- = 0. \end{aligned} \quad (1.3)$$

In dieser Form ist die linke Seite F in (1.3) ein Beispiel einer *verallgemeinerten Kojima-Funktion* (von \mathbb{R}^{n+m} in \mathbb{R}^{n+m}) im Sinne von [14, 13]. Sie wurde ursprünglich für nichtlineare Optimierungsaufgaben

$$\min\{f(x) \mid g_i(x) \leq 0, \quad i = 1, \dots, m\} \quad (1.4)$$

in [15] mit $p(x) = -Df(x)$ für eine bequeme Darstellung des Karush-Kuhn-Tucker (KKT-) Systems eingeführt.

Eine andere übliche Darstellung des Problems (1.1), (1.2) ist mittels einer sogenannten *NCP-Funktion* möglich, die als Funktion $G : \mathbb{R}^2 \rightarrow \mathbb{R}$ durch die Forderung

$$G^{-1}(0) = \{(s, t) \mid s \geq 0, t \geq 0, st = 0\} \quad (1.5)$$

definiert ist. Man untersucht dann anstelle von (1.3) das System

$$\begin{aligned} \Pi_1(x, y) &:= -p(x) + yDg(x) = 0 \\ \Pi_{2i}(x, y) &:= G(-g_i(x), y_i) = 0, \quad i = 1, \dots, m. \end{aligned} \quad (1.6)$$

Zahlreiche konkrete Ansätze für G findet man z.B. in [30, 5, 9].

Die Gleichungen (1.3) $F(x, y) = 0$ und (1.6) $\Pi(x, y) = 0$ haben zwar dieselben Lösungen x , sind aber strukturell unterschiedlich.

Ihnen ist gemeinsam, dass sie die Komplementaritätsbedingung im Originalproblem (1.1), (1.2) durch eine Gleichung beschreiben und so das Originalproblem selbst als ein System von $n + m$ Gleichungen mit $n + m$ reellen Variablen darstellen. In beiden Fällen sind die entsprechenden Vektorfunktionen F , Π in den Lösungen nicht notwendig differenzierbar.

Unterschiede bestehen im Zusammenhang mit Regularitätseigenschaften der Gleichungen, bei der Berechnungsmöglichkeit für diverse verallgemeinerte Ableitungen und bei daraus resultierenden Lösungsverfahren erster und zweiter Ordnung, speziell bei Verfahren vom Newton-Typ.

In der vorliegenden Arbeit sollen diese Unterschiede und Gemeinsamkeiten für Variationsaufgaben und Komplementaritätsprobleme aus einer einheitlichen Sicht untersucht und insbesondere die aus den verschiedenen Ansätzen resultierenden "nichtglatten" Newton-Verfahren mit herkömmlichen Verfahren vom Straf-, Barriere- und SQP-Typ verglichen werden.

Unser Ausgangspunkt ist eine (kleine) gezielte Störung des Kojima-Systems (1.3) mit $p(x) = -Df(x)$, mit der sich Ansätze vom Straf- und Barriere-Typ beschreiben lassen. Mit ihrer Hilfe wird gezeigt:

- Bei speziellen solcher Störungen fallen die Newton-Hilfsaufgaben für beide Gleichungen (1.3), (1.6) zusammen.
- Damit lösen Verfahren auf der Grundlage von NCP-Funktionen in jedem Schritt ein derart gestörtes Kojima-System.
- Darüberhinaus entsprechen allen diesen Ansätze im Newton-Verfahren einfach passende Bewichtungen λ_i (abhängig vom aktuellen Iterationspunkt) für die Gleichung

$$\lambda_i(g_i(x) + Dg_i(x)(\xi - x)) + (1 - \lambda_i)\eta_i = 0 \quad (1.7)$$

der linearisierten, komplementären Funktionen g_i und y_i .

- Ferner können, im Prinzip, diese Newton-Verfahren in einheitlicher Weise als SQP-Verfahren geschrieben werden (sequentielle quadratische Optimierung). "Im Prinzip" heißt hier, dass dann, wenn p in (1.1) keine Ableitung ist, die jeweilige quadratische Optimierungsaufgabe zu einem linearen Komplementaritätsproblem wird und man im Newton-Schritt ebenso wie im Falle $p = -Df$ eine lineare Gleichung zu lösen hat.

Die Interpretation als SQP-Verfahren hat den Vorteil, dass man direkt die Form der Hilfsprobleme (als quadratische Optimierungsaufgaben) kennt, die in jedem Schritt gelöst werden müssen. Das ermöglicht theoretisch ein besseres Verständnis der Methode und praktisch ein Umschalten zwischen Verfahren erster und zweiter Ordnung (z.B. wenn eine Newton-Matrix degeneriert oder der Startpunkt schlecht ist).

So wie die meisten Optimierungsverfahren stellt sich ein SQP-Verfahren nicht als eine einzige Methode dar, sondern als ein Zugang, von dem mehrere konkrete Algorithmen abgeleitet werden können. Die Hauptidee besteht darin, dass man das nichtlineare Problem, etwa (1.4), mittels einer Reihe von quadratischen Optimierungsaufgaben modelliert. Im Schritt k wird, ausgehend von einer approximativen Lösung x^k (zulässig oder nicht) für (1.4), ein quadratisches Modell von (1.4) um den Punkt x^k gelöst (zumeist auch mit Startpunkt x^k), um eine bessere Näherung x^{k+1} zu konstruieren.

Wir werden diese Methoden als Newton-Methode, angewandt auf die KKT-Bedingungen für Aufgabe (1.4) oder auf ein äquivalentes Gleichungssystem auffassen. Da die Funktionen F und Π nicht mehr überall differenzierbar sind, müssen hier allerdings spezielle Arten von verallgemeinerten Ableitungen zur Konstruktion der Hilfsaufgaben betrachtet werden, und die Wahl dieser Ableitungen bestimmt ebenfalls die quadratischen Ersatzprobleme.

Terminologie und Bezeichnungen orientieren sich in der vorliegenden Arbeit hauptsächlich an der üblichen Notation in der Optimierung und nichtglatten Analysis, wie sie etwa in [28], [14] benutzt werden. Mit einem Doppelpfeil $\Gamma : X \rightrightarrows Y$ wird eine mehrwertige Abbildung von X in Y bezeichnet, Summen der Form $A + rB$ sind für Teilmengen A, B desselben normierten Raumes über \mathbb{R} elementweise zu verstehen, $A + rB = \{x \mid x = a + rb, a \in A, b \in B\}$, und B (oder B_Y) bezeichnet die abgeschlossene Einheitskugel (im Raum Y). Im Falle $A = \{a\}$ schreibt man auch $a + rB$. Wir sagen, eine Aussage gilt für x nahe x^* , wenn sie für alle x in einer gewissen Umgebung von x^* richtig ist.

In der Arbeit, wenn nichts anderes gesagt wird, ist vorausgesetzt, dass alle Vektoren Spaltenvektoren sind. Außerdem habe die Jakobi Matrix von $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ die folgende Form:

$$Dg(x) = \begin{pmatrix} \frac{\partial g_1(x)}{\partial x_1} & \frac{\partial g_1(x)}{\partial x_2} & \dots & \frac{\partial g_1(x)}{\partial x_n} \\ \frac{\partial g_2(x)}{\partial x_1} & \frac{\partial g_2(x)}{\partial x_2} & \dots & \frac{\partial g_2(x)}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial g_m(x)}{\partial x_1} & \frac{\partial g_m(x)}{\partial x_2} & \dots & \frac{\partial g_m(x)}{\partial x_n} \end{pmatrix}$$

Kapitel 2

Verallgemeinerte Kojima–Funktion

2.1 Kojima–Funktion

Wir gehen aus von der Aufgabe

$$\min\{f(x) \mid h(x) = 0, g_i(x) \leq 0, i = 1, \dots, m\}, \quad (2.1)$$

die sich von (1.4) nur durch die zusätzliche Aufnahme von Gleichungsnebenbedingungen, beschrieben durch $h : \mathbb{R}^n \rightarrow \mathbb{R}^\kappa$ unterscheidet. Die Funktionen f, g, h seien vom Typ $C^{1,1}$, d.h. sie mögen lokal Lipschitz–stetige Ableitungen besitzen. Sei F komponentenweise definiert durch

$$\begin{aligned} F_1(x, y, z) &= Df(x) + \sum_{i=1}^m y_i^+ Dg_i(x) + \sum_{j=1}^\kappa z_j Dh_j(x) \\ F_{2,i}(x, y, z) &= g_i(x) - y_i^-, \quad i = \overline{1, m} \\ F_3(x, y, z) &= h(x) \\ \text{wobei } y_i^+ &= \max\{0, y_i\} \quad \text{und} \quad y_i^- = \min\{0, y_i\}. \end{aligned} \quad (2.2)$$

Die $C^{0,1}$ –Funktion $F : \mathbb{R}^{n+m+\kappa} \rightarrow \mathbb{R}^{n+m+\kappa}$ heißt dann *Kojima–Funktion* für Aufgabe (2.1).

Mit ihrer Hilfe kann man das KKT–System für (2.1) in einer äquivalenten Form als Gleichung schreiben

$$F(s) = 0 \quad \text{mit } s = (x, y, z) \quad \text{und} \quad F(s) = (F_1(s), F_2(s), F_3(s))^T, \quad (2.3)$$

wobei offenbar die folgenden Beziehungen gelten:

$$\begin{aligned} (x, y, z) \text{ KKT-Punkt} &\Rightarrow (x, y + g(x), z) \text{ erfüllt (2.3)} \\ (x, y, z) \text{ erfüllt (2.3)} &\Rightarrow (x, y^+, z) \text{ KKT-Punkt.} \end{aligned}$$

Es wurden hier einfach die Ungleichungen und die Komplementarität in den KKT–Bedingungen mit Hilfe der Funktionen y^+ und y^- ausgedrückt. Diese Lipschitz–Funktionen sind zwar nicht differenzierbar, haben aber eine sehr einfache Form.

Betrachten wir zunächst den wichtigen Fall $f, g, h \in C^2$. Dann gehört F zu Klasse der stückweisen C^1 –Funktionen (PC^1), die bereits intensiv im Rahmen nichtglatter Analysis untersucht worden ist. Insbesondere macht die Nullstellenbestimmung solcher Funktionen mittels Newton–Verfahren keine anderen Probleme als im klassischen C^1 –Fall, wenn man DF mit Hilfe der im Iterationspunkt aktiven C^1 –Funktionen definiert.

Gilt auch noch $y_i \neq 0 \forall i$, so existiert DF und ist außerdem stetig auf einer Umgebung von s , wonach das (lokale) Verhalten der Nullstellen bei kleinen Störungen von F durch den klassischen

Satz über implizite Funktionen beschrieben werden kann. Die Menge Θ^1 dieser C^1 -Punkte ist offenbar dicht in $\mathbb{R}^{n+m+\kappa}$, und für jedes $s = (x, y, z) \in \mathbb{R}^{n+m+\kappa}$ folgt

$$s(\varepsilon) := (x, y_1 + \varepsilon, \dots, y_m + \varepsilon, z) \in \Theta^1, \quad (2.4)$$

wenn nur $\varepsilon > 0$ hinreichend klein ist. Damit bietet sich die Möglichkeit an, überall eine verallgemeinerte Ableitung "D" F mittels

$${}^{\text{D}}F(s) = \lim_{\varepsilon \rightarrow +0} DF(s(\varepsilon)) \quad (2.5)$$

zu definieren und für Lösungsmethoden zu nutzen. Dasselbe bleibt mit

$${}^{\text{D}}F(s) = \lim_{\varepsilon \rightarrow -0} DF(s(\varepsilon)) \quad (2.6)$$

richtig. Allerdings können diese Typen einer "Ableitung" nicht zur Charakterisierung von Minima oder impliziter Funktionen benutzt werden, was man schon am Beispiel der Betragsfunktion sieht. In diesem Zusammenhang braucht man andere verallgemeinerte Ableitungen, die im nächsten Kapitel definiert werden.

Um solche Ableitungen nun auch im Falle $f, g, h \in C^{1,1}$ zu berechnen, schreiben wir F in Produktform. Dazu bilden wir den Zeilen-Vektor

$$N(y, z) := (1, y^+, y^-, z) \in \mathbb{R}^{1+2m+r},$$

der nur von den dualen Variablen abhängt, und eine Matrix, die nur die Originalfunktionen enthält:

$$M(x) := \begin{pmatrix} Df(x)^T & g(x)^T & h(x)^T \\ Dg(x) & 0 & 0 \\ 0 & -E & 0 \\ Dh(x) & 0 & 0 \end{pmatrix} \quad (2.7)$$

Die Matrix ist vom Typ $(1 + 2m + \kappa) \times (n + m + \kappa)$. Nun gilt

$$F(x, y, z) = N(y, z)M(x). \quad (2.8)$$

Diese Darstellung von F ist deshalb sehr hilfreich bei der Berechnung wichtiger verallgemeinerter Ableitungen, weil sie wie wir sehen werden der üblichen Produktregel genügen.

2.2 Verallgemeinerte Kojima–Funktion

Wenn wir die (Fréchet-) Ableitungen $Df(x)$, $Dg(x)$ und $Dh(x)$ in der ersten Spalte der Matrix $M(x)$ durch beliebige Funktionen $\Phi(x)$, $\Psi(x)$ und $\Gamma(x)$ von entsprechender Dimension ersetzen, so heißt

$$F^g = N(y, z)M^g(x)$$

verallgemeinerte Kojima–Funktion [13]. Hier bleiben die Komponenten F_2 und F_3 gegenüber (2.2) unverändert:

$$\begin{aligned} F_1^g &:= \Phi(x) + \sum_{i=1}^m y_i^+ \Psi_i(x) + \sum_{j=1}^r z_j \Gamma_j(x) \\ F_{2,i}^g &:= g_i(x) - y_i^-, \quad i = \overline{1, m} \\ F_3^g &:= h(x). \end{aligned} \quad (2.9)$$

Das System kann beispielweise auch folgendermaßen interpretiert werden. Sei $C = \{x \in \mathbb{R}^n \mid g(x) \leq 0, h(x) = 0\}$. Jeder Nullstelle $s = (x, y, z)$ von F^g entspricht nun eine Lösung x des Systems

$$x \in C, \quad -\Phi(x) \in \text{con} \{\Psi_i(x) \mid g_i(x) = 0\} + \text{lin} \cup_j \Gamma_j(x) \quad (2.10)$$

und umgekehrt ('con' bezeichnet die konische Hülle, 'lin' die lineare). Ist außerdem

$$\Psi_i(x) = Dg_i(x) \text{ und } \Gamma_j(x) = Dh_j(x) \quad \forall i, j \quad (2.11)$$

und erfüllt $x \in C$ eine Regularitätsbedingung, so bedeutet (2.10) gerade

$$0 \in \Phi(x) + N_C(x), \quad (2.12)$$

wenn $N_C(x)$ (wie im nichtkonvexen Fall üblich) als Polarkegel zum Kontingenzkegel (oder auch Bouligand–Kegel) [1] an C im Punkt x verstanden wird.

Die durch eine NCP-Funktion definierte Gleichung (1.6) nimmt für das System (2.9) die folgende Form an:

$$\begin{aligned} \Pi_1(x, y) &:= -\Phi(x) + \sum_i y_i \Psi_i(x) + \sum_j z_j \Gamma_j(x) = 0 \\ \Pi_{2i}(x, y) &:= G(-g_i(x), y_i) = 0, \quad i = \overline{1, m} \\ \Pi_3(x) &:= h(x) = 0. \end{aligned} \quad (2.13)$$

Die Verallgemeinerung der Kojima–Funktion ermöglicht es uns, mit denselben analytischen Mitteln wie für (gewöhnliche) Kojima–Funktionen auch andere Probleme mittels (2.9) zu untersuchen. Der Spezialfall (2.11) ist dabei von besonderer Bedeutung, weil er die Komponente F_1 mit der Beschreibung des zulässigen Bereiches in den Komponenten F_2, F_3 verbindet.

2.3 Beispiele

2.3.1 Komplementaritätsprobleme

Klassische Aufgabe Es seien u und $v : \mathbb{R}^n \rightarrow \mathbb{R}^n$ zwei gegebene Funktionen. Das Komplementaritätsproblem besteht darin, ein solches x zu finden, welches die Bedingungen

$$u(x) \geq 0, \quad v(x) \geq 0, \quad \langle u, v \rangle = 0 \quad (2.14)$$

erfüllt, wobei zumeist noch $v(x) = x$ gilt. Offenbar können wir die Aufgabe auch so umschreiben:

$$\begin{aligned} u(x) &= y^+ \\ v(x) &= -y^- \end{aligned}$$

wobei $y \in \mathbb{R}^n$ ist. Mit

$$F_1(x, y) := u(x) - y^+ \text{ und } F_2(x, y) := -v(x) - y^-$$

erhält man eine verallgemeinerte Kojima–Funktion $F(x, y) : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$. Hier sind $\Phi(x) := u(x)$, $\Psi(x) := -E$, $g(x) := -v(x)$, und z, F_3 entfallen. Die Bedingung (2.11) bedeutet hier gerade $v(x) = x$.

Ergänzte Aufgabe. Man kann z und F_3 benutzen, um eine zusätzliche Gleichung $w(x) = 0$ in das Modell (2.14) aufzunehmen. Letzteres ist oft dann von Interesse, wenn die Dimensionen des Bild- und Urbildraumes von $u, v : \mathbb{R}^n \rightarrow \mathbb{R}^m$ nicht identisch sind ($m < n$). Dies trifft zum Beispiel für unsere Ausgangs-Variationsaufgabe (1.1), (1.2) zu, wo man x in (2.14) aus dem primal-dual Vektor (x, y) bilden und

$$u(x, y) = y, \quad v(x, y) = -g(x), \quad w(x, y) = -p(x) + y^T Dg(x) \quad (2.15)$$

setzen kann. Dem Modell

$$u \geq 0, \quad v \geq 0, \quad w = 0, \quad \langle u, v \rangle = 0 \quad (2.16)$$

entspricht so die zu (1.1), (1.2) und damit auch zu (1.3) sowie (1.6) äquivalente Gleichung $F = 0$ mit den Komponenten

$$F_1 = u(x, y) - \lambda^+, \quad F_2 = -v(x, y) - \lambda^-, \quad F_3 = w(x, y) \quad (2.17)$$

Dies ist eine verallgemeinerte Kojima–Funktion, die die Bedingung (2.11) nicht erfüllt: $\Phi := u$, $\Psi := -E$, $g := -v$, $h = w$, $\Gamma = 0$.

2.3.2 Nash–Gleichgewicht im Stackelberg–Spiel

Seien $u, v : \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ und $u, v \in C^1$. Außerdem seien $X \in \mathbb{R}^n$ und $Y \in \mathbb{R}^m$ zwei nichtleere, kompakte, konvexe Mengen. Man finde $(x^0, y^0) \in X \times Y$, so dass

$$\begin{aligned} u(x^0, y^0) &\geq u(x, y^0) \quad \forall x \in X \\ v(x^0, y^0) &\geq v(x^0, y) \quad \forall y \in Y \end{aligned} \quad (2.18)$$

Wenn die Mengen X und Y durch (endlich viele) C^1 –Ungleichungen $p(x) \leq 0$ bzw. $q(y) \leq 0$ definiert sind, dann sehen die (unter einer Regularitätsbedingung) notwendigen Optimalitätsbedingungen für (2.18) so aus:

$$\begin{aligned} -D_x u(x, y) + r^+ Dp(x) &= 0 \\ -D_y v(x, y) + s^+ Dq(y) &= 0 \\ p(x) - r^- &= 0, \quad q(y) - s^- = 0. \end{aligned} \quad (2.19)$$

wobei r, s die entsprechenden dualen Vektoren sind. Erneut hat (2.19) die Form einer verallgemeinerten Kojima–Funktion, die (2.11) erfüllt. Hier sind

$$\Phi(x, y) = (-D_x u(x, y), -D_y v(x, y))^T \quad \text{und} \quad \Psi(x, y) = (Dp(x), Dq(y))^T,$$

d.h. die Matrix M^g hängt jetzt von $x' := (x, y)$ ab und der Vektor N von $y' = (r, s)$.

Beinhaltet das Spiel noch gemeinsame Bedingungen für die Strategien beider Spieler

$$g(x, y) \leq 0, \quad (2.20)$$

(solche Spiele, auch von B. Rosen [29] untersucht, heißen oft Stackelberg–Spiele), so nimmt (2.19) mit dazu gehörenden Lagrange–Vektoren λ_1, λ_2 für die beiden Spieler die Form

$$\begin{aligned} -D_x u(x, y) + r^+ Dp(x) + \lambda_1^+ D_x g(x, y) &= 0 \\ -D_y v(x, y) + s^+ Dq(y) + \lambda_2^+ D_y g(x, y) &= 0 \\ p(x) - r^- &= 0, \quad q(y) - s^- = 0, \\ g(x, y) &= \lambda_1^-, \quad g(x, y) = \lambda_2^- \end{aligned} \quad (2.21)$$

an. Das ist kein verallgemeinertes Kojima-System mit der Eigenschaft (2.11), weil in den beiden ersten Gleichungen die Terme $\lambda_1^+ D_x g(x, y)$ beziehungsweise $\lambda_2^+ D_y g(x, y)$ fehlen. Zum anderen kann man nicht einfach generell $\lambda_1 = \lambda_2$ verlangen, was erneut zu (2.11) führen würde.

Man kann auch Beispielen ohne Beschränkung auf kompaktes X bzw. Y begegnen. Sei $u = 0$, $v = y$, $g = x + y \in \mathbb{R}$, wonach (2.21) die Form

$$\begin{aligned} 0 + \lambda_1^+ &= 0, & -1 + \lambda_2^+ &= 0 \\ x + y &= \lambda_1^-, & x + y &= \lambda_2^- \end{aligned}$$

annimmt.

Ist allerdings der zulässige Bereich $M = X \times Y$ konvex, kompakt und nichtleer, und sind die C^1 -Funktionen u, v in x bzw. y konkav, so kann man jedem Paar $(x^0, y^0) \in M$ die Lösungsmenge $\Lambda(x^0, y^0)$ der Aufgabe

$$\max_{(x, y) \in M} u(x, y^0) + v(x^0, y)$$

zuordnen (dieser Trick ist der Kern der Arbeit von Rosen [29]). Sie ist ebenfalls konvex, kompakt und nichtleer, und $(x, y) \in \Lambda(x^0, y^0)$ impliziert unter einer Regularitätsbedingung (Abadie-Bedingung [19]) die Existenz von λ mit

$$\begin{aligned} -D_x u(x, y^0) + r^+ D p(x) + \lambda^+ D_x g(x, y) &= 0 \\ -D_y v(x^0, y) + s^+ D q(y) + \lambda^+ D_y g(x, y) &= 0 \\ p(x) - r^- &= 0, \quad q(y) - s^- = 0, \quad g(x, y) = \lambda^-. \end{aligned} \tag{2.22}$$

Weiter besitzt die (abgeschlossene, mehrwertige) Abbildung $\Lambda : M \rightrightarrows M$ nach Kakutani einen Fixpunkt. Mit $(x, y) = (x^0, y^0) \in \Lambda(x, y)$ erhält man dann aber einen Gleichgewichtspunkt, der (2.21) tatsächlich mit $\lambda_1 = \lambda_2$ erfüllt, und dieser Punkt wird erneut durch ein verallgemeinertes Kojima-System mit (2.11) beschrieben. Diese Aussagen lassen sich direkt auf Gleichgewichtspunkte in Spielen mit drei und mehr Spielern übertragen.

2.3.3 Eine verallgemeinerte Gleichung

In Analogie zur Variationsungleichung (1.1) suche man ein $x \in C$, so dass

$$p(x) \in H(x), \tag{2.23}$$

wobei H durch eine Funktion w gegeben ist

$$H(x) = \{\xi \in \mathbb{R}^n \mid w(x, \xi, t) \leq 0 \ \forall t \in T\}.$$

Hier müssen die Menge $T \neq \emptyset$ konvex, kompakt und die Funktion w in t konkav sein. Weiter sei C wieder durch $g(x) \leq 0$ definiert und T durch $\tau(t) \leq 0$. Alle auftretenden Funktionen seien stetig differenzierbar. Das Supremum von w bzgl. t wird in einem t^* angenommen, und unter Regularität der Nebenbedingungen folgt

$$D_t w(x, \xi, t^*) + y^+ D \tau(t^*) = 0, \quad \tau(t^*) = y^-. \tag{2.24}$$

Damit wird aus dem Originalproblem (2.23) das System

$$\begin{aligned} D_t w(x, p(x), t) + y^+ D \tau(t) &= 0, \\ \tau(t) = y^-, \quad g(x) = \lambda^-, \quad w(x, p(x), t) &= \mu^-. \end{aligned} \tag{2.25}$$

Es ist mittels einer verallgemeinerten Kojima–Funktion F mit Primalvariablen (x, t) und dem Dualvektor (y, λ, μ) darstellbar, ohne (2.11) zu erfüllen: $\Phi = D_t w(x, p(x), t)$, die Ψ –Komponenten zu τ werden aus $D\tau(t)$ gebildet, die zu den restlichen Ungleichungen sind Null.

Kapitel 3

Ableitungen und Regularität

3.1 Regularität und $TF, CF, \partial F$

Seien X und Y zwei normierte Räumen, $S : Y \rightrightarrows X$ eine mehrwertige Abbildung und $(y^*, x^*) \in \text{gph } S$.

Definition 3.1 Die Abbildung S heißt *pseudo-Lipschitz (mit Rang L) im Punkt (y^*, x^*)* , falls es Umgebungen U von x^* und V von y^* gibt, so dass für jeden Punkt $(y, x) \in (V, U) \cap \text{gph } S$ und $y' \in V$ ein $x' \in S(y')$ mit

$$\|x' - x\|_X \leq L \|y' - y\|_Y$$

existiert.

Diese Eigenschaft wird auch oft als Aubin-property bezeichnet [28].

Definition 3.2 Die Abbildung S heißt *upper-Lipschitz (mit Rang L) im Punkt (y^*, x^*)* , falls es Umgebungen U von x^* und V von y^* gibt, so dass für jeden Punkt $(y', x') \in (V, U) \cap \text{gph } S$ ein $x \in S(y^*)$ mit

$$\|x' - x\|_X \leq L \|y' - y^*\|_Y$$

existiert.

Sei $S = F^{-1}$ die inverse Abbildung einer Multifunktion $F : X \rightrightarrows Y$.

Definition 3.3 Wenn S pseudo-Lipschitz im Punkt (y^*, x^*) ist, so heißt F *pseudo regulär in (x^*, y^*)* . Falls darüber hinaus solche U und V existieren, dass $U \cap F^{-1}(y)$ einelementig ist für alle $y \in V$, so nennt man F *streng regulär in (x^*, y^*)* . Ist S upper-Lipschitz in (y^*, x^*) , so nennen wir F *upper regulär in (x^*, y^*)* .

In dieser Arbeit wird außerdem die folgende Verschärfung von Definition 3.3 für stetige Funktionen betrachtet.

Definition 3.4 ([16]): Eine stetige Funktion $F : X \rightarrow Y$ heißt *streng regulär im Punkt $s^* \in X$ bezüglich einer Untermenge $G \subset C^{0,1}(X, Y)$* , wenn es reelle $\theta > 0$ und $\rho > 0$ gibt, so dass die Gleichung

$$F(s) + \varphi(s) = F(s^*) \tag{3.1}$$

für $s \in B(s^*, \rho)$ und $\varphi \in G$ eine eindeutige und Lipschitz-stetige Lösung $s = s(\varphi)$ besitzt, sofern $\|\varphi\|_{B(s^*, \rho)}^{0,1} < \theta$.

Im Falle $\varphi \in C^1$ ist dies der Begriff der strengen Regularität aus [26].

Die Norm $\|\varphi\|_{\Omega}^{0,1}$ ist für eine beschränkte Menge $\Omega \subset X$ wie folgt erklärt:

$$\|\varphi\|_{\Omega}^{0,1} = \max\{\sup_{s \in \Omega} \|\varphi(s)\|, \text{Lip}(\varphi(s)/\Omega)\},$$

und $\text{Lip}(\varphi(s)/\Omega)$ steht für die kleinste Lipschitz-Konstante von φ auf Ω .

In analoger Weise, nämlich durch Ersetzen der Umgebung V in Definition 3.1 durch eine Kugel um die konstante Funktion $\phi(s) = y^*$, kann man pseudo-Lipschitz, pseudo regulär und upper regulär auf die nichtlineare Variation von F verallgemeinern (natürlich wird dann auf die lokale Einzigkeit der Lösung von (3.1) verzichtet). Einfache Variationen ϕ dieser Art (stückweise linear) werden unsere gestörten Kojima-Funktionen definieren.

Die angegebenen Regularitätstypen lassen sich mit Hilfe verallgemeinerter Ableitungen charakterisieren. Sei dazu $z = (x, y) \in \text{gph } F$ und $F : X \rightrightarrows Y$.

Definition 3.5 $CF(z) : X \rightrightarrows Y$ heißt Contingent-Ableitung an der Stelle z in Richtung u , wenn $v \in CF(z)(u) \Leftrightarrow \exists t \downarrow 0$ und $(u_t, v_t) \rightarrow (u, v)$, so dass $y + tv_t \in F(x + tu_t)$.

Insbesondere, wenn $F : X \rightarrow Y$ eine Funktion ist, liegt $y = F(x)$ fest, und es folgt

$$CF(x)(u) = \left\{ v \mid v = \lim_{t \downarrow 0} \lim_{u_t \rightarrow u} \frac{F(x + tu_t) - F(x)}{t}, \text{ für gewisse Folgen} \right\}.$$

Definition 3.6 $TF(z) : X \rightrightarrows Y$ heißt Thibault-Ableitung an der Stelle z in Richtung u , wenn $v \in TF(z)(u) \Leftrightarrow \exists t \downarrow 0, (x_t, y_t) \rightarrow (x, y)$ in $\text{gph } F$ und $(u_t, v_t) \rightarrow (u, v)$, so dass $y_t + tv_t \in F(x_t + tu_t)$.

Insbesondere, wenn $F : X \rightarrow Y$ eine Funktion ist, folgt

$$TF(x)(u) = \left\{ v \mid v = \lim_{t \downarrow 0} \lim_{(x_t, u_t) \rightarrow (x, u)} \frac{F(x_t + tu_t) - F(x_t)}{t}, \text{ für gewisse Folgen} \right\}.$$

Eine besondere Rolle spielen hier lokal Lipschitz-stetige Funktionen F (kurz $F \in C^{0,1}$) in endlicher Dimension, weil dann die Bildmengen $CF(x)(u)$ und $TF(x)(u)$ nichtleer sind und die konvergenten Folgen $u_t \rightarrow u$ durch konstante $u_t = u$ ersetzt werden können, was die Berechnung der Limites vereinfacht. Ist dann F in x richtungsdifferenzierbar, d.h. existiert

$$F'(x; u) = \lim_{t \rightarrow +0} \frac{F(x + tu) - F(x)}{t},$$

so folgt auch $CF(x)(u) = \{F'(x; u)\}$. Außerdem beschreiben TF und CF die eingeführten Regularitätsbegriffe vollständig.

Satz 3.1 Sei $F \in C^{0,1}(\mathbb{R}^n, \mathbb{R}^n)$. Dann ist F streng regulär an der Stelle x genau dann, wenn $TF(x)$ injektiv ist, d.h. $0 \notin TF(x)(\mathbb{R}^n \setminus \{0\})$. In diesem Fall ist F auch streng regulär in x bezüglich $G = C^{0,1}(\mathbb{R}^n, \mathbb{R}^n)$.

Für einen Beweis von Satz 3.1 siehe [16].

Satz 3.2 Sei $F \in C^{0,1}(\mathbb{R}^n, \mathbb{R}^m)$. Dann ist F upper regulär an der Stelle x genau dann, wenn $CF(x)$ injektiv ist, d.h. $0 \notin CF(x)(\mathbb{R}^n \setminus \{0\})$. In diesem Fall ist F auch upper regulär in x bezüglich $G = C^{0,1}(\mathbb{R}^n, \mathbb{R}^n)$.

Einen Beweis von Satz 3.2 findet man in [12] oder auch in [14]. Die upper Regularität bezüglich $C^{0,1}$ folgt hier elementar, weil $F^{-1}(y')$ leer sein darf (siehe Definition 3.3), d.h. Lösbarkeit von (3.1) wird nicht verlangt.

Satz 3.3 Sei $F \in C^{0,1}(\mathbb{R}^n, \mathbb{R}^m)$. Dann ist F pseudo regulär an der Stelle x genau dann, wenn $CF(x)$ uniform surjektiv ist, d.h. es gibt ein $\varepsilon > 0$, so dass

$$\varepsilon B_m \subset CF(x')(B_n), \quad \forall x' \in x + \varepsilon B_n.$$

In diesem Fall ist F auch pseudo regulär in x bezüglich $G = C^{0,1}(\mathbb{R}^n, \mathbb{R}^m)$.

Einen Beweis von Satz 3.3 findet man in [18]. Für entsprechende Aussagen in allgemeinen Räumen und mittels Ko–Ableitungen siehe [23] und [14].

Die hier betrachteten Regularitätsbegriffe kennzeichnen jeweils eine bestimmte (lokale) Stabilität der Lösungen des Gleichungssystems (3.1). Auf der anderen Seite treten üblicherweise Regularitätsforderungen auf, wenn man die Originalgleichung mittels eines (verallgemeinerten) Newton–Verfahrens lösen will (uniform reguläre Newton–Matrizen für Iterationspunkte nahe der Lösung, Injektivität der benutzten Ableitung nahe der Lösung), siehe Bedingung (BI) in Kapitel "Newton–Verfahren" (Seite 25). Für das gewöhnliche Newton–Verfahren und C^1 –Funktionen $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ bedeutet diese Forderung gerade $\det DF(x^*) \neq 0$ und damit strenge Regularität im Sinne obiger Definition.

Für $C^{0,1}$ –Funktionen, die wir hier hauptsächlich betrachten, hängt diese Forderung (ebenso wie die Schwierigkeit der Newton–Hilfsaufgabe) von der benutzten Art einer verallgemeinerten Ableitung ab. Wir müssen deshalb das Newton–Verfahren und die entsprechende "Newton–Regularität" für $C^{0,1}$ –Funktionen etwas genauer betrachten und beginnen damit, mögliche Ableitungen vorzustellen. Sie haben nicht direkt mit klassischen Stabilitätseigenschaften der Gleichung (3.1) zu tun.

3.2 Ableitungen für Newton–Verfahren

3.2.1 Generalized Jacobian

Sei $F \in C^{0,1}(\mathbb{R}^n, \mathbb{R}^m)$. Wir definieren eine Menge

$$\Theta(F) := \{x \in \mathbb{R}^n \mid \text{die Fréchet–Ableitung von } f \text{ existiert in } x\}.$$

Nach dem Satz von Rademacher hat die Komplementärmenge $\mathbb{R}^n \setminus \Theta(F)$ das Lebesgue–Maß Null. Außerdem ist die Norm von $DF(x')$ für $x' \in \Theta(F)$ und x' nahe einem festen x durch einen lokalen Lipschitz–Rang L von F begrenzt. Dies garantiert uns, dass die Abbildung $\partial_0 F : \mathbb{R}^n \rightrightarrows \mathbb{R}^{m \times n}$ mit

$$\partial_0 F(x) := \{A \mid A = \lim_{\substack{x' \rightarrow x \\ x' \in \Theta(F)}} DF(x'), \quad \text{für gewisse Folgen}\}$$

nichtleere Bilder besitzt. Man kann auch leicht sehen, dass $\partial_0 F$ (als mehrwertige Abbildung) abgeschlossen und lokal kompakt ist. Dasselbe gilt dann auch für die folgende Abbildung ∂F :

Definition 3.7 (Clarke 1976 [2]) Die konvexe Hülle von $\partial_0 F(x)$

$$\partial F(x) := \text{conv } \partial_0 F(x),$$

heißt (Clarke's) generalized Jacobian.

Die Menge $\partial_0 F(x)$ wird auch oft B-Differential genannt und mit $\partial_B F(x)$ bezeichnet. Die grundlegende Motivation dieser Definitionen wird durch Clarke's Inversensatz [2] geliefert.

Satz 3.4 Sei $F \in C^{0,1}(\mathbb{R}^n, \mathbb{R}^n)$ und enthalte $\partial F(x)$ keine singuläre Matrix (d.h. die Abbildung $u \rightarrow \partial F(x)u$ ist injektiv). Dann ist F streng regulär an der Stelle x .

Im Gegensatz zu Satz 3.1 ist die Nicht-Singularität von $\partial F(x)$ keine notwendige Bedingung für die strenge Regularität (siehe Beispiel in [16]).

3.2.2 Limes von Jacobians von C^1 -Punkten

Analog werden in [14] die Menge

$$\Theta^1(F) := \{x \in \mathbb{R}^n \mid F \text{ ist } C^1 \text{ auf einer Umgebung von } x\}$$

und die Ableitung

$$D^0 F(x) = \{A \mid A = \lim_{x' \rightarrow x, x' \in \Theta^1(F)} DF(x'), \text{ für gewisse Folgen}\}$$

definiert. Allerdings können $\Theta^1(F)$ und $D^0 F(x)$ leer sein.

Definition 3.8 Eine $C^{0,1}$ -Funktion F heißt pseudo-glatt, wenn $\Theta^1(F)$ dicht im Definitionsbereich ist.

Diese Eigenschaft besitzen viele $C^{0,1}$ -Funktionen in konkreten Anwendungen. Aus den Definitionen folgt unmittelbar

$$D^0 F(x) \subset \partial_0 F(x) \subset \partial F(x) \text{ und } CF(x)(u) \subset TF(x)(u).$$

Außerdem gilt der folgende Satz [14]:

Satz 3.5 Es gilt

- (i) $\partial_0 F(x)u \subset TF(x)(u)$
- (ii) $TF(x)(u) \subset \partial F(x)u$ und
- (iii) $\partial F(x)u = \text{conv}(TF(x)(u))$.

Beweis:

Sei $A \in \partial_0 F(x)$. Nach der Definition gibt es eine Folge $x^k \rightarrow x$ mit $x^k \in \Theta(F)$ und $DF(x^k) \rightarrow A$. Dann findet man (hinreichend kleine) $t_k \rightarrow +0$ so, dass gilt

$$F(x^k + t_k u) - F(x^k) = t_k DF(x^k)u + o(t_k)$$

Das bedeutet

$$\lim_{k \rightarrow \infty} DF(x^k)u = Au \in TF(x)(u)$$

Damit haben wir die erste Inklusion gezeigt. Wenn wir nun die Menge $\partial_0 F(x)u$ durch ihre konvexe Hülle ersetzen, folgt

$$\partial F(x)u \subset \text{conv } TF(x)(u). \quad (3.2)$$

Um die Behauptung (ii) zu beweisen, kann man den Mittelwertsatz [3] anwenden

$$F(x^k + t_k u) - F(x^k) \in \text{conv } [\cup_{0 \leq \theta \leq 1} \partial F(x^k + \theta t_k u)(t_k u)].$$

Da ∂F als mehrwertige Abbildung abgeschlossen und lokal beschränkt ist, erhalten wir mit $x^k \rightarrow x$ und $t_k \rightarrow +0$ die Inklusion (ii). Diese Inklusion zusammen mit (3.2) zeigt nun, dass auch (iii) wahr ist. ■

3.3 Anwendungen auf die Kojima–Funktion

Im folgenden werden zentrale Aussagen von [14] benutzt.

3.3.1 Die Ableitungen $\partial F, TF$ und CF

Der Grund für die besondere Rolle von (verallgemeinerten) Kojima Funktionen liegt darin, dass die Ableitungen TF und CF den "normalen" Differentiationsregeln genügen und deshalb (relativ) leicht berechnet werden können.

Satz 3.6 [14] *Die Ableitungen TF und CF einer verallgemeinerten Kojima–Funktion $F = M(x)N(y, z)$ können, für $g, h, \Phi, \Psi, \Gamma \in C^{0,1}$, mittels ihrer partiellen Ableitungen bestimmt werden, d.h.,*

$$TF(x, y, z)(u, v, w) = T_x F(x, y, z)(u) + T_{(y,z)} F(x, y, z)(v, w), \quad (3.3)$$

was wegen der Produktform dasselbe sagt wie

$$TF(x, y, z)(u, v, w) = TM(x)(u)N(y, z) + M(x)TN(y, z)(v, w). \quad (3.4)$$

Die Formeln (3.3), (3.4) gelten auch für die Contingent–Ableitung. Sie bleiben richtig, wenn $M = M(x)$ eine beliebige $C^{0,1}$ –Matrixfunktion entsprechender Dimension ist und $N = N(y, z)$ komponentenweise eine Linearkombination von y_i^+, y_i^- und z darstellt.

Da F unter unseren Voraussetzungen immer lokal Lipschitz-stetig ist, stimmt CF mit der Richtungsableitung F' genau dann überein, wenn die Bildmengen $CF(s)$ stets einelementig sind.

C^1 –Funktion M

Ist $M = M(x)$ eine C^1 –Funktion, so unterscheiden sich $TM(x) = CM(x) = DM(x)$ nicht, und es spielen nur die unterschiedlichen Ableitungen von N eine Rolle. Wegen der leicht zu sehenden Identität

$$TN(y, z)(v, w) = \partial N(y, z)(v, w)$$

folgt dann auch

$$TF(x, y, z)(v, w) = \partial F(x, y, z)(v, w).$$

Wir betrachten jetzt den Fall einer Kojima–Funktion (2.2) mit $f, g_i, h_j \in C^2$ gesondert. In jeder dieser Aussage kann man auch Df durch eine C^1 –Funktion $p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ersetzen (wie in (2.11)). Wir sprechen dann immer noch vom " C^2 –Fall".

Dann ist Injektivität von TF in $s = (x, y, z)$ äquivalent zur Regularität der generalized Jacobian $\partial F(s)$, und das bedeutet wiederum Regularität aller Matrizen der folgenden Form

$$\mathcal{M}(s, r) := \begin{pmatrix} D_x F_1(s) & r_1 Dg_1(x) & \cdots & r_m Dg_m(x) & Dh(x) \\ Dg_1(x)^T & -(1 - r_1) & & & 0 \\ \vdots & 0 & \ddots & 0 & \vdots \\ Dg_m(x)^T & & & -(1 - r_m) & 0 \\ Dh(x)^T & 0 & \cdots & 0 & 0 \end{pmatrix} \quad (3.5)$$

wobei r in der Menge

$$R_T(y) = \left\{ r \in \mathbb{R}^m \mid r_i = \begin{cases} 1, & y_i > 0 \\ [0, 1], & y_i = 0 \\ 0, & y_i < 0 \end{cases}, i = 1, 2, \dots, m \right\} \quad (3.6)$$

liegen soll.

Die Menge $R_T(y)$ besteht komponentenweise aus den Thibault–Ableitung $T(y_i^+)(y_i)(v_i)$ der Funktionen $y_i \rightarrow y_i^+$ im Punkt y_i in Richtung $v_i = 1$. Anstelle der Thibault–Ableitung kann man auch der generalized Jacobian nehmen, da sie hier, wie es schon erwähnt wurde, übereinstimmen

$$T(y^+)(y)(v) = \partial(y^+)(y)(v) \quad \forall y, v \in \mathbb{R}^m.$$

Weiter gilt für $r \in R_T(y)$ der folgende Satz [8], [14].

Satz 3.7 *Alle Matrizen $\mathcal{M}(s, r)$ (3.5) mit $r \in R_T(y)$ sind regulär \Leftrightarrow alle Determinanten*

$$\det \mathcal{M}(s, r) \text{ mit } r_i \in \{0, 1\} \text{ und } r \in R_T(y)$$

haben dasselbe Vorzeichen und sind verschieden von Null.

Explizit erhält man schliesslich mit $s = (x, y, z)$:

Lemma 3.1 *Im Falle $f, g, h \in C^2$, besteht die Menge $TF(s)(u, v, w)$ aus allen Produkten*

$$\mathcal{M}(s, r)(u, v, w)^T \text{ mit } r \in R_T(y);$$

die Menge $CF(s)(u, v, w)$ besteht aus denselben Produkten, nur ist zusätzlich

$$r_i = \begin{cases} 1, & \text{falls } y_i = 0 \text{ und } v_i > 0 \\ 0, & \text{falls } y_i = 0 \text{ und } v_i \leq 0 \end{cases}$$

zu erfüllen (kurz $r \in R_C(y, v)$).

Mit der Variablentransformation

$$\alpha_i = r_i v_i \text{ und } \beta_i = (1 - r_i) v_i \quad (3.7)$$

bedeutet dann die Inklusion

$$P \in TF(s)(u, v, w) \quad (3.8)$$

das Bestehn der folgenden Gleichungen:

$$\begin{array}{rcccccc}
 P_1 = & D_x F_1(s)u & +\alpha_1 Dg_1(x) & \cdots & +\alpha_m Dg_m(x) & +Dh(x)w \\
 P_{2,1} = & Dg_1(x)^T u & -\beta_1 & & & \\
 & \vdots & & & & \\
 P_{2,m} = & Dg_m(x)^T u & & & & -\beta_m \\
 P_3 = & Dh(x)^T u & 0 & \cdots & & 0
 \end{array} \tag{3.9}$$

wobei $r \in R_T(y)$ hier äquivalent ist zu

$$\begin{array}{l}
 \alpha_i = 0 \text{ wenn } y_i < 0; \beta_i = 0 \text{ wenn } y_i > 0 \\
 \text{und } \alpha_i \beta_i \geq 0 \text{ wenn } y_i = 0.
 \end{array} \tag{3.10}$$

Analog bedeutet die Inklusion

$$P \in CF(s)(u, v, w) \tag{3.11}$$

wieder die lineare Gleichung (3.9), nur ist jetzt $r \in R_C(y, v)$ äquivalent zu

$$\begin{array}{l}
 \alpha_i = 0 \text{ wenn } y_i < 0; \beta_i = 0 \text{ wenn } y_i > 0 \\
 \text{und } \alpha_i \beta_i = 0, \alpha_i \geq 0, \beta_i \geq 0 \text{ wenn } y_i = 0.
 \end{array} \tag{3.12}$$

Die Inklusion (3.8) kann man also auch durch lineare Ungleichungssysteme beschreiben, die Inklusion (3.11) durch ein lineares Komplementaritätsproblem, das von (3.9) und (3.12) gegeben ist.

Die Gestalt der Ableitungen bleibt wegen der Produktregel unverändert, wenn man verallgemeinerte Kojima–Funktionen mit $\Phi, \Psi, \Gamma \in C^1$ betrachtet. Entsprechend (2.9) hat dann F_1 die Form

$$F_1^g := \Phi(x) + \sum_{i=1}^m y_i^+ \Psi_i(x) + \sum_{j=1}^r z_j \Gamma_j(x),$$

und es ist nur jeweils Df, Dg_i und Dh durch Φ, Ψ_i und Γ in der ersten Zeile von (3.9) und (3.5) zu ersetzen.

3.3.2 Die Ableitungen $\partial_0 F$ und $D^0 F$

Wieder betrachten wir zuerst den C^2 –Fall. Da die Abbildung $M = M(x)$ aus der Produktdarstellung (2.8) von F stetig differenzierbar ist, ist F nur dann nicht C^1 auf einer hinreichend kleinen Umgebung von $s = (x, y, z)$, wenn eine Komponente y_i Null ist. Die Mengen $\Theta(F)$ und $\Theta^1(F)$ fallen hier also zusammen, und so werden beide Ableitungen $\partial_0 F$ und $D^0 F$, indem man für die Folgen in der Definition

$$y'_i = y_i + \varepsilon \text{ und } y'_i = y_i - \varepsilon \tag{3.13}$$

setzt, nun durch alle Matrizen der Form (3.5) mit speziellerem r beschrieben:

$$D^0 F(s) = \partial_0 F(s) = \{M(s, r) \mid r \in R_T(y) \text{ und } r_i \in \{0, 1\} \forall i\}. \tag{3.14}$$

Ist die Matrixfunktion $M = M(x)$ aus (2.8) nur vom Typ $C^{0,1}$, dann geht die Übereinstimmung verloren, und es kann sogar $\Theta^1(M) = \emptyset$ sein (Beispiele BE 0 und BE 1 in [14]).

Kapitel 4

Newton–Verfahren

Im folgenden wollen wir unterschiedliche verallgemeinerte Ableitungen für das Newton–Verfahren untersuchen und nutzen. Deshalb gehen wir von einem allgemeinen Modell aus, das in [17] erstmals behandelt wurde.

4.1 Allgemeine Konvergenzaussagen

Sei eine Funktion $F : X \rightarrow Y$ (normierte Räume) in der Nähe von $x^* \in X$ Lipschitz–stetig mit Konstante L und sei $F(x^*) = 0$. Weiter sei

$$GF : (X, X) \rightrightarrows Y$$

eine mehrdeutige Abbildung mit

$$GF(x, u) \neq \emptyset \quad \text{und} \quad GF(x, 0) = \{0\}. \quad (4.1)$$

Wir wollen die Abbildung $GF(x^k, \cdot)$ als irgendeine verallgemeinerte Richtungsableitung der Funktion F an der Stelle x^k ansehen und mit ihrer Hilfe ein Newton–Verfahren formulieren.

Zu gegebenem x^k wollen wir daher eine Lösung u für

$$0 \in F(x^k) + GF(x^k, u) \quad (4.2)$$

bestimmen und

$$x^{k+1} = x^k + u \quad (4.3)$$

setzen. Wir lassen dabei zu, dass die Inklusion (4.2) nur mit einer Genauigkeit $\alpha > 0$ erfüllt wird, d.h. wir begnügen uns mit einem u , das die Bedingung

$$\alpha \|F(x^k)\|_B \cap [F(x^k) + GF(x^k, u)] \neq \emptyset \quad (4.4)$$

erfüllt (ohne uns für die Lösbarkeit von (4.2) zu interessieren).

Definition 4.1 *Das Tripel (F, GF, x^*) heißt zulässig, wenn für jedes $\varepsilon \in (0, 1)$ positive r und α existieren, so dass der Prozess (4.3), (4.4) mit Anfangspunkt $x^0 \in x^* + rB$ immer lösbar bleibt und für die Iterationspunkte die folgende Konvergenzrelation gilt*

$$\|x^{k+1} - x^*\| \leq \varepsilon \|x^k - x^*\|.$$

Die beiden nächsten Sätze zeigen, dass sich die Zulässigkeit von (F, GF, x^*) im Wesentlichen auf die folgenden Bedingungen reduziert.

Bedingung (BI) (Injektivität der Ableitung nahe x^*)

$$\|v\| \geq c\|u\|, \quad \forall v \in GF(x, u), \quad \forall u \in X \quad \forall x \text{ nahe } x^* \quad (c > 0 \text{ und fest}).$$

Bedingung (BA) (Approximation)

$$F(x) - F(x^*) + GF(x, u) \subset GF(x, x + u - x^*) + o(x - x^*)B \quad \forall u \in X.$$

Wenn uns hier nur die Richtung $u = x^* - x$ interessiert, erhalten wir wegen $GF(x, 0) = \{0\}$ eine schwächere Bedingung:

Bedingung (BA)* (vereinfachte (BA))

$$F(x) - F(x^*) + GF(x, x^* - x) \subset o(x - x^*)B.$$

Lemma 4.1 *Die Bedingungen (BA) und (BA)* sind äquivalent, wenn X und Y endliche Dimension besitzen und $GF(x, u)$ irgendeine der hier betrachteten Ableitungen ist; also wenn $GF(x, u)$ zusammenfällt mit $TF(x)(u)$, $CF(x)(u)$, $\partial F(x)u$, $\partial_0 F(x)u$, $D^0 F(x)u$ oder (im Falle der Richtungs-differenzierbarkeit) mit $F'(x; u)$.*

Beweis:

Für $TF(x)(u)$, $CF(x)(u)$, $F'(x; u)$ siehe [17]. Für die generalized Jacobian $\partial F(x)u$ besteht $GF(x, x^* - x)$ in (BA)* aus Elementen der Form $A(x^* - x)$, $A \in \partial F(x)$. Mit derselben Matrix A folgt dann wegen Linearität aus

$$F(x) - F(x^*) + A(x^* - x) \in o(x - x^*)B$$

auch

$$F(x) - F(x^*) + Au \in A(x + u - x^*) + o(x - x^*)B.$$

Derselbe Schluß, nur mit anderen Mengen an der Stelle von $\partial F(x)$, zeigt die Behauptung für $\partial_0 F$ und $D^0 F$. ■

Der nächste Satz formuliert eine hinreichende Bedingung für die Konvergenz des Newton–Verfahrens (4.4), der nachfolgende liefert notwendige Bedingungen. Die Beweise findet man in [17] und [14].

Satz 4.1 *Das Tripel (F, GF, x^*) ist zulässig, wenn eine Konstante $c > 0$ existiert, so dass die Bedingungen (BI) und (BA) erfüllt sind.*

Ist die Bedingung (BI) erfüllt, sprechen wir auch von *Newton–Regularität der Ableitung* GF nahe x^* . Erfüllt eine Funktion $F \in C^{0,1}(\mathbb{R}^n, \mathbb{R}^m)$ die Bedingung (BA)* mit $GF(x, u) = \partial F(x)u$, so heißt sie auch *semismooth* in x^* . Der Begriff stammt von Mifflin [22] und wurde in den Arbeiten [24], [25] und anderen aufgegriffen.

Es werden jetzt die Konvergenzbedingungen des Newton Verfahrens verbunden mit den Regularitätsbedingungen aus §3.1 für die Gleichung (3.1).

Satz 4.2 Wir nehmen an, dass $F \in C^{0,1}(\mathbb{R}^n, \mathbb{R}^n)$.

1) Sei $GF(x, u) = TF(x)(u)$.

Es gilt: (BI) ist erfüllt $\Leftrightarrow F$ ist streng regulär in $(x^*, 0)$. Wenn (BI) gilt, dann ist die Bedingung (BA) notwendig für die Zulässigkeit von (F, GF, x^*) .

2) Sei $GF(x, u) = \partial F(x)u$.

Es gilt: (BI) ist erfüllt $\Leftrightarrow \partial F(x^*)$ enthält keine singuläre Matrizen. Diese Bedingung ist stärker als strenge Regularität.

3) Mögen die Richtungsableitung $F'(x; u)$ nahe x^* existieren und gelte $GF(x, u) = F'(x; u)$.

3.1) Unter strenger Regularität von F an der Stelle $(x^*, 0)$ ist (BA) notwendig und hinreichend für die Zulässigkeit von (F, F', x^*) .

3.2) Unter pseudo Regularität von F an der Stelle $(x^*, 0)$ ist (BI) erfüllt.

Die letzte Aussage 3.2) dieses Satzes ist eine Folgerung aus dem Satz von P. Fusek [6], [7], der unter anderem für eine richtungsdifferenzierbare und in $(x^*, 0)$ pseudo reguläre Funktion $F \in C^{0,1}(\mathbb{R}^n, \mathbb{R}^n)$ gerade die Eigenschaft (BI) sichert. Insbesondere muß dann x^* eine isolierte Nullstelle von F sein.

4.2 Approximation von $GF(x, u)$ und $F(x)$

Man kann $GF(x, u)$ durch eine andere Abbildung $\hat{G}F(x, u) : (X, X) \rightrightarrows Y$ ersetzen, ohne die Bedingungen (BI) und (BA) zu verletzen, solange gilt

$$\emptyset \neq \hat{G}F(x, u) \subset GF(x, u) + O(x - x^*)\|u\|B. \quad (4.5)$$

Um das zu zeigen betrachten wir ein Element $\hat{v}(x, u) \in \hat{G}F(x, u)$. Aus (4.5) erhalten wir dann die folgende Darstellung:

$$\hat{v}(x, u) = v(x, u) + \delta(x)\|u\|, \quad (4.6)$$

wobei $v(x, u) \in GF(x, u)$ und $\|\delta(x)\| \leq O(x - x^*)$ gilt.

(BI) für $\hat{G}F(x, u)$. Erfülle $GF(x, u)$ die Bedingung (BI). Es folgt dann :

$$\|\hat{v}\| \geq \|v\| - \|\delta\|\|u\| \geq c\|u\| - \|\delta\|\|u\| = (c - \|\delta\|)\|u\|, \forall u \in X$$

Die Abbildung $\hat{G}F(x, u)$ erfüllt damit die Bedingung (BI) mit einer Konstante $\hat{c} = c - \|\delta(x)\| > 0$, sobald x genug nahe an x^* liegt.

(BA) für $\hat{G}F(x, u)$. Erfülle $GF(x, u)$ die Bedingung (BA). D.h. für beliebiges $v(x, u) \in GF(x, u)$ existieren ein Element $v(x, x + u - x^*) \in GF(x, x + u - x^*)$ und $r(x)$ mit $\|r(x)\| \leq o_1(x - x^*)$, so dass die folgende Gleichung gilt.

$$F(x) - F(x^*) + v(x, u) = v(x, x + u - x^*) + r(x)$$

Sie ist aber äquivalent zur Gleichung

$$F(x) - F(x^*) + v(x, u) + \delta(x)\|u\| = v(x, x + u - x^*) + \delta(x)\|x + u - x^*\| \\ + r(x) + \delta(x)(\|u\| - \|x + u - x^*\|),$$

wobei $\delta(x)$ ein beliebiges Element mit $\|\delta(x)\| \leq O(x - x^*)$ ist. Sei nun

$$\hat{r}(x) = r(x) + \delta(x)(\|u\| - \|x + u - x^*\|).$$

Mit (4.6) erhalten wir dann

$$F(x) - F(x^*) + \hat{v}(x, u) = \hat{v}(x, x + u - x^*) + \hat{r}(x). \quad (4.7)$$

Die Norm von der Funktion $\hat{r}(x)$ schätzt sich folgendermaßen ab

$$\|\hat{r}(x)\| \leq \|r(x)\| + \|\delta(x)\|(\|x - x^*\| + \|u\| - \|u\|) \leq o_2(x - x^*).$$

Da $\hat{v}(x, u) \in \hat{G}F(x, u)$ in der Gleichung (4.7) beliebig ist, haben wir damit auch die Bedingung (BA) für die Abbildung $\hat{G}F(x, u)$ bewiesen.

Nun können wir auch zeigen, dass die Approximation der Ausgangsfunktion $F(x)$ durch eine andere $\hat{F}(x)$ mit der Eigenschaft:

$$\|\hat{F}(x) - F(x)\| \leq o(x - x^*) \quad (4.8)$$

ebenfalls keinen Einfluss auf die Bedingungen (BI) und (BA) besitzt.

Satz 4.3 *Das Tripel (F, GF, x^*) ist zulässig genau dann, wenn das Tripel $(\hat{F}, \hat{G}F, x^*)$ mit (4.8) und (4.5) zulässig ist, sobald $\hat{F}(x)$ die Injektivitätsbedingung $\|\hat{F}(x)\| \geq c\|x - x^*\|$, mit $c > 0$ erfüllt.*

Beweis:

Die Äquivalenz (im Sinne des Satzes) zwischen den Tripeln (F, GF, x^*) und $(F, \hat{G}F, x^*)$ folgt aus den Überlegungen oben und dem Satz 4.1. Es bleibt nur zu zeigen, dass die Tripel $(F, \hat{G}F, x^*)$ und $(\hat{F}, \hat{G}F, x^*)$ auch äquivalent sind. Tatsächlich, sei nach (4.8)

$$F(x) = \hat{F}(x) + \delta(x), \quad \text{wobei } \|\delta(x)\| \leq o(x - x^*)$$

und sei die Bedingung (4.4) für $(F, \hat{G}F, x^*)$ erfüllt. D.h. es existiert ein Element $\hat{v}(x, u) \in \hat{G}F(x, u)$, für das die folgende Relation gilt

$$\|\hat{F}(x) + \delta(x) + \hat{v}(x, u)\| \leq \alpha\|\hat{F}(x) + \delta(x)\|$$

Diese ist äquivalent zur Ungleichung

$$\|\hat{F}(x) + \delta(x) + \hat{v}(x, u)\| + \|-\delta(x)\| \leq \alpha\|\hat{F}(x) + \delta(x)\| + \|\delta(x)\|.$$

Was letztendlich liefert:

$$\|\hat{F}(x) + \hat{v}(x, u)\| \leq \alpha\|\hat{F}(x)\| + (1 + \alpha)\|\delta(x)\|.$$

Da die Funktion $\hat{F}(x)$ injektiv ist, gilt insbesondere

$$\|\hat{F}(x)\| \geq \|\delta(x)\|.$$

Somit erhalten wir

$$\|\hat{F}(x) + \hat{v}(x, u)\| \leq \beta \|\hat{F}(x)\|.$$

D.h. die Bedingung (4.4) für $(\hat{F}, \hat{G}F, x^*)$ mit einer Konstante $\beta = 2\alpha + l > \alpha$ ist auch erfüllt. Ausgehend von der Definition 4.1 haben wir damit den Satz bewiesen. ■

Bemerkung 4.1 Wenn das Tripel (F, GF, x^*) die Bedingungen (BI) und (BA)* erfüllt, erhält man die benötigte Injektivitätsbedingung für $\hat{F}(x)$. Tatsächlich, nach (BA)* existiert also ein Element $v(x, u) \in GF(x, x - x^*)$, so dass die Gleichung

$$F(x) = -v(x, u) + r_1(x)$$

mit $\|r_1(x)\| \leq o(x - x^*)$ gilt. Der Approximation (4.8) nach erhalten wir gleich auch

$$\hat{F}(x) = -v(x, u) + r_2(x),$$

wobei $\|r_2(x)\| \leq o(x - x^*)$. Für $v(x, u)$ gilt die Bedingung (BI). D.h.

$$\|\hat{F}(x)\| = \|-v(x, u) + r_2(x)\| \geq \|v(x, u)\| - \|r_2(x)\| \geq c\|x - x^*\|$$

q.e.d.

4.3 Die Anwendung von $D^0F, LocPC^1$ und $F'(x; u)$

Sei weiter $F \in C^{0,1}(\mathbb{R}^n, \mathbb{R}^n)$.

4.3.1 D^0F und $LocPC^1$

Erfüllt eine Abbildung GF die Bedingungen (4.1), (BI) und (BA)*, so offenbar auch jede (Auswahl-) Abbildung Γ mit

$$\emptyset \neq \Gamma(x; u) \subset GF(x; u).$$

Andererseits wollen wir in C^1 -Punkten x^k von F kein neues Newton-Verfahren erfinden. Deshalb definieren wir jetzt

$$GF(x; u) = DF(x)u \quad \forall x \in \Theta^1(F). \quad (4.9)$$

Falls F , wie wir nun voraussetzen, pseudo-glatt ist, liegt damit GF auf der offenen und dichten Teilmenge $\Theta^1(F)$ des Definitionsbereiches fest. Betrachtet man nun die *kleinste abgeschlossene Abbildung GF mit (4.9)* (" GF abgeschlossen" bedeutet, dass der Graph von GF eine abgeschlossene Menge ist), so besteht $GF(x; u)$ aus allen Limites der Form

$$g = \lim DF(x')u' \text{ mit } x' \in \Theta^1(F), u' \rightarrow u \text{ und } x' \rightarrow x.$$

Weil weiter DF lokal beschränkt bleibt (pseudo-glatt fordert insbesondere $F \in C^{0,1}$), folgt hieraus, dass

$$GF(x; u) = D^0F(x)u \quad (4.10)$$

gilt. Die Ableitung D^0F hat somit eine Minimaleigenschaft in Bezug auf das Newton-Verfahren. Wir nehmen nun an, dass $RF(x) \in D^0F(x), \forall x \in \mathbb{R}^n$ gilt.

Lemma 4.2 *Wenn F pseudo–glatt ist und $(BA)^*$ für eine Auswahlfunktion RF von D^0F erfüllt ist, so ist die Bedingung für alle Auswahlfunktionen von D^0F erfüllt, und es gilt*

$$CF(x^*)(u) \subset D^0F(x^*)u \quad \forall u.$$

Beweis:

Sei $SF \in D^0F$ eine Auswahlfunktion und $RF \neq SF$. Man stellt fest, $RF(x) = SF(x)$, $\forall x \in \Theta^1$, denn es gilt

$$D^0F(x) = DF(x), \quad \forall x \in \Theta^1(F), \quad (4.11)$$

d.h. SF erfüllt $(BA)^*$ in Punkten $x^* + u' \in \Theta^1$. Wenn die Funktion $o(\cdot)$ in $(BA)^*$ nicht oberhalbstetig ist, können wir sie immer durch solch eine ersetzen (z.B. $o_{sup}(u) = \limsup_{u' \rightarrow u} o(u') \geq o(u)$). Da F pseudo–glatt ist, kann man jeden Punkt $x^* + u$ als Limes ($u' \rightarrow u$) von Elementen $x^* + u'$ aus Θ^1 darstellen. Dank der Stetigkeit gilt somit, dass die Bedingung $(BA)^*$ für SF in jedem Punkt erfüllt bleibt.

Um die Inklusion zu beweisen, betrachtet man den folgenden Quotient:

$$b(t) = t^{-1}[F(x^* + tu(t)) - F(x^*)]$$

Durch die pseudo–Glattheit von F approximiert $b(t)$ (mit Fehler $< t$) den Quotient:

$$a(t) = t^{-1}[F(x^* + tu) - F(x^*)]$$

so dass $\|u(t) - u\| < t$ und $x^* + tu(t) \in \Theta^1$ gilt. Nach Definition 3.5 besteht die Menge $CF(x^*)(u)$ aus Limites genau von solchen $a(t)$ mit gewissen Folgen $t \rightarrow +0$. Mit $(BA)^*$ und (4.11) folgt außerdem

$$b(t) \in DF(x^* + tu(t))u(t) + t^{-1}o(tu(t))B.$$

Wir erhalten also $\lim a(t) = \lim b(t) \in D^0F(x^*)u$, wenn $t \rightarrow +0$. ■

Bemerkung 4.2 Die Existenz einer Auswahlfunktion RF von D^0F , die $(BA)^*$ erfüllt, ist eine wesentliche Voraussetzung. Ohne sie gilt die Behauptung des Lemmas nicht. Man kann nämlich ein Gegenbeispiel finden (siehe [14]), wo $(BA)^*$ nicht erfüllt ist, obwohl $DF(x^*)$ existiert. Auch die reelle Funktion $F(x) = x^2 \sin(1/x)$ ($x \neq 0$), $F(0) = 0$ zeigt diesen Effekt.

Die folgende Funktionenklasse $LocPC^1$ wird in [14] etwas allgemeiner definiert. Im folgenden reicht uns eine speziellere, dafür aber übersichtlichere Definition.

Definition 4.2 *Eine pseudo–glatte Funktion $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ gehört zu Klasse $LocPC^1$, wenn eine endliche Überdeckung des \mathbb{R}^n durch konvexe Polyeder $\{P^s | s = 1, \dots, N\}$ derart existiert, dass F auf $\text{int } P^s$ stetig differenzierbar ist und, für jedes kompakte $K \subset \mathbb{R}^n$, DF gleichmäßig stetig auf $\text{int } P^s \cap K$ ist.*

Insbesondere gehören hierzu alle NCP –Funktionen, die in Abschnitt 6.1 studiert werden. Der folgende Satz vermittelt uns die Begründung für die Betrachtung solcher Funktionen.

Satz 4.4 Wenn $F(x) \in LocPC^1$ gilt, so erfüllt $GF(x, u) = D^0F(x)u$ die Bedingung $(BA)^*$.

Beweis:

Sei $x^* \in R^n$ gegeben und sei F auf Θ^1 nach wie vor eine C^1 -Funktion. Wir betrachten einen Punkt $y = x^* + u \in \Theta^1 \cap (x^* + rB)$. Nach der Definition 4.2 existiert ein konvexes Polyeder P^s , so dass gilt

$$\text{int conv } \{x^*, y\} \subset \text{int } P^s,$$

sofern r genügend klein ist. Außerdem wissen wir, dass $F \in C^1$ auf $\text{int } P^s$ ist. Das erlaubt uns folgendermaßen abzuschätzen:

$$\begin{aligned} F(y) - F(x^*) &= DF(x^* + \theta u)u = \\ &= DF(x^* + \theta u)u + DF(x^* + u)u - DF(x^* + u)u, \end{aligned}$$

wobei $\theta \in (0, 1)$. Daraus erhält man sofort

$$F(y) - F(x^*) - DF(x^* + u)u \in \|u\| \sup_{\theta \in (0,1)} \|DF(x^* + \theta u) - DF(x^* + u)\|B.$$

Da DF hier gleichmäßig stetig ist, kann man das Supremum durch $O(u)$ nach oben abschätzen. Es gilt also

$$F(x^* + u) - F(x^*) - DF(x^* + u)u \in o(u)B, \quad (4.12)$$

für alle $x^* + u \in \Theta^1 \cap (x^* + rB)$. Betrachten wir nun alle mögliche Folgen $u \rightarrow u^*$, so dass $x^* + u \in \Theta^1 \cap (x^* + rB)$ gilt. Für unterschiedliche Folgen erhält man i.a. unterschiedliche Grenzwerte in der linken Seite von (4.12). Da F stetig ist, die Funktion $o(u)$ o.B.d.A. als oberhalb stetig angesehen werden kann und DF beschränkt bleibt, bleibt die Inklusion (4.12) bei Grenzübergang bestehen. Aus der Definition von D^0F folgt daher, dass für alle u^* gilt:

$$F(x^* + u^*) - F(x^*) - D^0F(x^* + u^*)u^* \subset o(u^*)B,$$

d.h. die Bedingung $(BA)^*$ ist an der Stelle x^* mit $GF(x, u) = D^0F(x)u$ erfüllt. ■

4.3.2 Verknüpfte Funktionen

In den folgenden Anwendungen ist F (wie z.B. eine Kojima-Funktion) eine verknüpfte Funktion vom Typ $F(x) = h(z(x))$. Sind hier h und z pseudo-glatt, braucht das nicht mehr für F zu stimmen, weil z.B. z in einen Unterraum abbilden kann, der gerade nicht zu $\Theta^1(h)$ gehört. Analog könnte die Eigenschaft $LocPC^1$ verloren gehen und D^0F gar nicht definiert sein. Man kann deshalb nicht erwarten, dass D^0F allgemein eine Kettenregel in Gleichungsform wie in Satz 3.6 erfüllt. Allerdings kann man die Abbildung

$$W(x, u) = D^0h(z(x))[D^0z(x)u] := \{D^0h(z(x))v \mid v \in [D^0z(x)u]\} \quad (4.13)$$

bilden und für das Newton-Verfahren benutzen.

Lemma 4.3 Mit $h, z \in LocPC^1$ folgt, dass $GF = W$ weiter die Bedingung $(BA)^*$ erfüllt.

Beweis:

Man kann direkt die Abschätzungen im letzten Beweis nun auf h und z und Punkte in $\Theta^1(h)$, $\Theta^1(z)$ anwenden und zum Limes in Bezug auf nun zwei Folgen übergehen. ■

In [14] wird eine analoge Aussage mit Hilfe der Verknüpfung von "Newton maps" bewiesen, die wir hier nicht betrachten.

Für eine Kojima–Funktion folgt so im C^2 –Fall (auch mit $p = Df$ wie in §3.3): $D^0F(s)$ besteht aus allen Matrizen aus (3.14).

Unter einem NCP–Zugang (siehe Abschnitt 6.1), der das Problem (2.14)

$$u(x) \geq 0, \quad v(x) \geq 0, \quad \langle u, v \rangle = 0$$

als eine Gleichung

$$G(z_i(x)) = 0; \quad z_i = (u_i, v_i); \quad i := 1, 2, \dots, n$$

löst, kann man die Ableitung $W(x, u) = D^0G(z(x))[D^0z(x)u]$ nutzen, sofern G und $z = (u, v)$ zu $LocPC^1$ gehören.

Es bleibt dann allein das Problem, die Newton–Regularität, d.h., die Bedingung (BI) zu untersuchen.

Falls $F(x)$ richtungsdifferenzierbar ist, stimmt $CF(x)(u)$ mit $F'(x; u)$ überein, d.h. Lemma 4.1, Lemma 4.2, Satz 4.4, Satz 4.1 und die Aussage 3.2) von P. Fusek im Satz 4.2 liefern uns den

Satz 4.5 *Ist $F \in LocPC^1$, pseudo regulär an der Stelle $(x^*, 0)$ und richtungsdifferenzierbar in allen Punkten einer Umgebung von x^* , so sind die Bedingungen (BI) und (BA) mit $GF(x, u) = F'(x; u)$ erfüllt, d.h. (F, F', x^*) ist zulässig.*

Anders gesagt, unter solchen Voraussetzungen wird das Newton–Verfahren mit Richtungsableitungen $F'(x; u)$ funktionieren. Das können wir aber nicht behaupten, wenn wir mit der Thibault–Ableitung $TF(x)(u)$ bzw. mit Clarke's generalized Jacobian $\partial F(x)$ arbeiten. Nach dem Satz 4.2 braucht man hier nämlich strenge Regularität von F in $(x^*, 0)$ bzw. Regularität aller Matrizen aus $\partial F(x^*)$. Diese Bedingungen sind im allgemeinen stärker als die Forderung nach pseudo-Regulärheit. Für eine Kojima–Funktion fallen sie jedoch im C^2 –Fall (auch mit $p = Df$ wie in §3.3) nach dem Hauptergebnis von Donchev und Rockafellar in [4] zusammen.

Man muß allerdings hier zugeben, dass die Newton–Gleichungen mit gewöhnlichen Richtungsableitungen $F'(x; u)$ allgemein sehr schwer zu lösen sind. In den folgenden Kapiteln werden wir deshalb das Newton Verfahren mit anderen verallgemeinerten Richtungsableitungen betrachten. Z.B. für die Kojima–Systeme rechnen wir Thiboult–Ableitung bzw. generalized Jacobian aus. Diese Ableitungen sind durch Satz 3.6 gegeben, spezieller durch Lemma 3.1. Während für NCP–Funktionen D^0 –Ableitung vorgezogen wird. Das liegt einfach daran, dass die Thibault–Ableitung $TG(u_i, v_i)(\cdot, \cdot)$ sowie die generalized Jacobian $\partial G(u_i, v_i)$ nicht in Gleichungsform bekannt sind, und andererseits in C^1 Punkten ohnehin mit der Fréchet–Ableitung gearbeitet werden soll.

Kapitel 5

Gezielte Störungen von Kojima-Funktionen

Die Anwendung eines Newton-Verfahrens auf die Kojima-Funktion (2.2), d.h. auf

$$\begin{aligned} F_1(x, y, z) &= Df(x) + \sum_{i=1}^m y_i^+ Dg_i(x) + \sum_{j=1}^{\kappa} z_j Dh_j(x) \\ F_{2,i}(x, y, z) &= g_i(x) - y_i^-, \quad i = \overline{1, m} \\ F_{3,j}(x, y, z) &= h_j(x), \quad j = \overline{1, \kappa} \end{aligned} \quad (5.1)$$

wobei $x \in \mathbb{R}^n$, $y_i^+ = \max\{0, y_i\}$ und $y_i^- = \min\{0, y_i\}$

macht nach dem vorigen Kapitel (lokal) keine Probleme, wenn die beschreibenden Funktionen vom Typ C^2 sind und strenge Regularität in der Nullstelle $s^* = (x^*, y^*, z^*)$ vorliegt. Dasselbe gilt für Variationsungleichungen, d.h., wenn Df durch eine C^1 -Funktion p entsprechender Dimension ersetzt wird.

Allerdings kann man (bei schlechtem Startpunkt) leicht zu unlösbaren Hilfsaufgaben gelangen: Um das zu sehen, nehme man an, dass $n < m$ gilt und im aktuellen Iterationspunkt $s = (x, y, z)$ die Komponenten y_1, \dots, y_m alle positiv sind. Die linearisierte 2. Gleichung des Kojima-Systems verlangt dann wegen $Dy_i^- = 0$, ein $u \in \mathbb{R}^n$ zu finden mit

$$g_i(x) + Dg_i(x)u = 0, \quad i = \overline{1, m}$$

Dieses Teilsystem ist aber überbestimmt.

Uns interessiert deshalb jetzt die Frage: Was passiert mit der Lösung von (2.3), wenn wir die Kojima-Funktion F ein bißchen verändern? Diese Veränderungen kann man als Störung oder Regularisierung der Originalfunktion auffassen. Unser Ausgangspunkt ist die Funktion F^ε mit den Komponenten

$$\begin{aligned} F_1(x, y, z) &:= Df(x) + \sum_{i=1}^m y_i^+ Dg_i(x) + \sum_{j=1}^{\kappa} z_j Dh_j(x) \\ F_{2,i}^\varepsilon(x, y, z) &:= g_i(x) - y_i^- - \varepsilon_i y_i^+, \quad i = \overline{1, m} \\ F_{3,j}^\varepsilon(x, y, z) &:= h_j(x) - \varepsilon_j z_j, \quad j = \overline{1, \kappa} \end{aligned} \quad (5.2)$$

Wenn wir das variierte System $F^\varepsilon = 0$ anschauen, können wir gleich seine wichtigste Eigenschaft sehen. Für $\varepsilon_j \neq 0$, hat das Teilsystem

$$F_{3,j}^\varepsilon(x, y, z) = 0, \quad j = \overline{1, \kappa}$$

(und seine Linearisierung) nun immer eine Lösung, unabhängig davon wie groß die Zahl κ ist. Das Gleiche gilt auch für

$$F_{2,i}^\varepsilon(x, y, z) = 0, \quad i = \overline{1, m},$$

falls $\varepsilon_i > 0$ ist, denn dann bildet die Funktion $y_i \rightarrow y_i^- + \varepsilon_i y_i^+$ auf ganz \mathbb{R} ab. Der obige Nachteil des Originalsystem (2.3) wird also kompensiert.

Im Folgenden betrachten wir die Gleichung $F^\varepsilon(s) = 0$ genauer, um zu verstehen, welche Rolle die Nullstellen dieser Gleichung spielen. Für die Störungen bzgl. der Gleichungen setzen wir stets $\varepsilon_j > 0$ voraus und untersuchen als erstes die verschiedenen Fälle für ε_i .

5.1 Beziehung zu den bekannten Verfahren

Quadratische Straffunktion

Fall 1. $\varepsilon_i > 0, \forall i$

Sei (x, y, z) eine Nullstelle von (5.2). Aus der 2. Gleichung von (5.2) folgt

$$\begin{aligned} \text{Wenn } y_i \geq 0, \text{ so folgt } & g_i(x) = \varepsilon_i y_i^+ \quad \text{d.h.} \quad y_i^+ = g_i^+(x)/\varepsilon_i. \\ \text{Wenn } y_i < 0, \text{ so folgt } & y_i^+ = 0 \quad \text{d.h.} \quad g_i(x)^+ = 0. \end{aligned}$$

Wir erhalten damit in jedem Falle die Gleichung

$$F_1(s) = Df(x) + \sum_{i=1}^m \frac{1}{\varepsilon_i} g_i^+(x) Dg_i(x) + \sum_{j=1}^{\kappa} \frac{1}{\varepsilon_j} h_j(x) Dh_j(x) = 0.$$

Sie bedeutet, dass x ein stationärer Punkt der Straffunktion

$$H_\varepsilon^S(x) := f(x) + \frac{1}{2} \sum_{i=1}^m \frac{1}{\varepsilon_i} g_i^+(x)^2 + \frac{1}{2} \sum_{j=1}^{\kappa} \frac{1}{\varepsilon_j} h_j(x)^2$$

ist. In der anderen Richtung gilt: Wenn x ein stationärer Punkt von $H_\varepsilon^S(x)$ ist, so wird (x, y, z) mit

$$z_j = \frac{h_j(x)}{\varepsilon_j} \quad \text{und} \quad y_i = \begin{cases} g_i(x)/\varepsilon_i, & \text{falls } g_i(x) > 0 \\ g_i(x), & \text{falls } g_i(x) \leq 0 \end{cases}$$

eine Nullstelle von (5.2).

Die Philosophie zur Strafmethode besteht darin, dass man, anstatt (2.1) zu lösen, eine Folge (bezüglich $\varepsilon \rightarrow +0$) freier Minimierungsprobleme

$$\lim_{\varepsilon \rightarrow +0} \min H_\varepsilon^S(x) \tag{5.3}$$

löst. Wenn $\varepsilon > 0$ hinreichend klein ist, schlägt sich eine verletzte Restriktion $g_i(x) > 0$ bzw. $h_j(x) \neq 0$ stark in der Zielfunktion H_ε^S nieder. Wird $H_\varepsilon^S(\cdot)$ minimiert über \mathbb{R}^n , sollten bei kleinem ε die Restriktionen wenigstens ungefähr erfüllt sein. Dazu gilt der Satz [21].

Satz 5.1 *Unter den Voraussetzungen*

(i) f, g_i seien konvex, alle h_j seien affin-linear

(ii) Der zulässige Bereich der Aufgabe (2.1) sei nichtleer und beschränkt

gilt: Ist ε hinreichend klein, so existieren Minimalpunkte von $H_\varepsilon^S(\cdot)$ über \mathbb{R}^n . Jede Folge $x(\varepsilon)$ von Minimalpunkten hat für $\varepsilon \rightarrow +0$ einen Häufungspunkt x^ . Jeder solche Häufungspunkt löst (2.1).*

Bei Anwendung der Methode treten oft verschiedene Probleme auf. Zum Beispiel: Wie schnell soll ε verkleinert werden, damit die letzte Lösung $x(\varepsilon)$ ein guter Startpunkt für die Suche nach $x(\varepsilon')$, $\varepsilon' < \varepsilon$ bleibt? Auch das Stopkriterium ist nicht offensichtlich. Hierzu gibt es keine allgemeinen Rezepte, solange man keine zusätzliche Information über $\|x(\varepsilon) - x(\varepsilon')\|$ besitzt.

Quadratische Barrierefunktion

Fall 2. $\varepsilon_i < 0, \forall i$

Sei (x, y, z) eine Nullstelle von (5.2). Aus der zweiten Gleichung in (5.2) folgt dann sofort, dass $g_i(x) \leq 0, \forall i$. Das bedeutet offenbar $g_i(x) = g_i^-(x), \forall i$. Im folgenden betrachten wir die Indexmenge

$$I := \{i \mid y_i > 0, i = \overline{1, m}\}. \quad (5.4)$$

Aus der 2. Gleichung von (5.2) haben wir dann:

$$\begin{array}{ll} \text{Wenn } i \in I, \text{ so folgt} & g_i^-(x) = \varepsilon_i y_i^+ \quad \text{d.h.} \quad 0 < y_i^+ = g_i^-(x)/\varepsilon_i. \\ \text{Wenn } i \notin I, \text{ so folgt} & y_i^+ = 0 \quad \text{d.h.} \quad g_i^-(x) = y_i^-. \end{array}$$

Die erste Gleichung in (5.2) sieht dann nach dem Einsatz von den anderen so aus

$$F_1(s) = Df(x) + \sum_{i \in I} \frac{1}{\varepsilon_i} g_i^-(x) Dg_i(x) + \sum_{j=1}^{\kappa} \frac{1}{\varepsilon_j} h_j(x) Dh_j(x) = 0. \quad (5.5)$$

Damit ist der Punkt x ein stationärer Punkt für die folgende Funktion.

$$H_{\varepsilon, I}^B(x) := f(x) + \frac{1}{2} \sum_{i \in I} \frac{1}{\varepsilon_i} g_i^-(x)^2 + \frac{1}{2} \sum_{j=1}^{\kappa} \frac{1}{\varepsilon_j} h_j(x)^2.$$

Wir werden im nächsten Abschnitt sehen, dass sie (nach einer Transformation) auch als Barrierefunktion verstanden werden kann.

Die Umkehrung gilt auch: Wenn wir einen stationären Punkt x von $H_{\varepsilon, I}^B(x)$ mit einer vorgegebenen Indexmenge I und mit $g_i(x) < 0, \forall i \in I$ sowie $g(x) \leq 0$ haben, dann löst der Punkt (x, y, z) mit der Eigenschaft

$$z_j = \frac{h_j(x)}{\varepsilon_j} \quad \text{und} \quad y_i = \begin{cases} g_i(x)/\varepsilon_i, & \text{falls } i \in I \\ g_i(x), & \text{falls } i \notin I \end{cases}$$

die Gleichung $F^\varepsilon(s) = 0$.

Die entsprechende Barriermethode hat dann die Form:

$$\lim_{\varepsilon \rightarrow -0} \max H_{\varepsilon, I}^B(x) \quad (5.6)$$

Es ist bemerkenswert, dass zumindest bei nur einer konvexen Ungleichung $g < 0$ und linearem f , der stationäre Punkt von $H_{\varepsilon, I}^B$ nahe x^* ein *Maximalpunkt* von $H_{\varepsilon, I}^B(x)$ ist. Dies ist leicht am Beispiel $\min\{x \mid x^2 \leq 1\}$ zu sehen.

Logarithmische Barrierefunktion

Man kann die Barrierefunktion $H_{\varepsilon, I}^B(x)$ weiter transformieren. Wir haben $g_i(x) = g_i^-(x) < 0$, $\forall i \in I$. Für solche i gilt also auch

$$g_i(x)^- Dg_i(x) = g_i(x)^2 D(\ln(-g_i(x))).$$

Das läßt uns die Gleichung (5.5) so umschreiben

$$F_1(s) = Df(x) + \sum_{i \in I} \frac{1}{\varepsilon_i} g_i(x)^2 D(\ln(-g_i(x))) + \sum_{j=1}^{\kappa} \frac{1}{\varepsilon_j} h_j(x) Dh_j(x) = 0.$$

Mit

$$\delta_i := -g_i(x)^2 / \varepsilon_i > 0 \text{ und } \delta_j := \varepsilon_j > 0 \quad (5.7)$$

folgt so

$$F_1(s) = Df(x) - \sum_{i \in I} \delta_i D(\ln(-g_i(x))) + \sum_{j=1}^{\kappa} \frac{1}{\delta_j} h_j(x) Dh_j(x) = 0.$$

Anderes gesagt, der Punkt x ist ein stationärer Punkt für die folgende gewöhnliche logarithmische Barrierefunktion (mit Penalty-Term für die Gleichungen)

$$H_{\delta, I}^L(x) := f(x) - \sum_{i \in I} \delta_i \ln(-g_i(x)) + \frac{1}{2} \sum_{j=1}^{\kappa} \frac{1}{\delta_j} h_j(x)^2.$$

Man muß hier allerdings beachten, dass der "Barrierefaktor" δ_i von x abhängt. Die klassische Barriermethode verlangt für den Barrierefaktor $\delta_i \rightarrow +0$. Deshalb sollte gelten (damit man von einer Barriermethode sprechen kann)

$$\delta_i \xrightarrow{\varepsilon_i \rightarrow 0} +0 \quad (5.8)$$

Später (im Abschnitt Abschätzungen) wird gezeigt, wie dies für Lösungen nahe x^* aus der strengen Regularität folgt. Das gestörte Kojima–System liefert uns dann ein entsprechendes logarithmisiertes Barriere-Verfahren

$$\lim_{\delta \rightarrow +0} \min H_{\delta, I}^L(x) \quad (5.9)$$

Aufgrund der Konvexität von $H_{\delta, I}^L$ auf der Slaterpunktmenge sieht man, dass wir hier (bei konvexen Problemen mit linearem h) wieder zur Suche globaler *Minima* zurückgekommen sind.

Umgekehrt, ein stationärer Punkt x von $H_{\delta, I}^L(x)$ mit einer vorgegebenen Indexmenge I und mit $g(x) \leq 0$ sowie $g_i(x) < 0 \forall i \in I$ liefert uns nach Definition von

$$z_j = \frac{h_j(x)}{\delta_j} \quad \text{und} \quad y_i = \begin{cases} -\delta_i / g_i(x), & \text{falls } i \in I \\ g_i(x), & \text{falls } i \notin I \end{cases}$$

eine Nullstelle (x, y, z) von (5.2).

Die Idee der Barriermethode besteht darin (wenn man von den Gleichungsrestriktionen absieht, die wieder als Strafterm in die Hilfs-Zielfunktion eingehen), dass man, statt (2.1) zu lösen, eine Folge (bezüglich $\delta = \delta(k)$) freier Minimierungsprobleme (5.9) behandelt. Der Startpunkt muß

im Inneren des zulässigen Bereiches liegen (genauer in der Nullstellenmenge von h und der Slaterpunktmenge von g). Wenn $\delta > 0$ hinreichend klein ist, darf der entsprechende Iterationspunkt immer näher zum Rand der Restriktionsmenge kommen. Er kann ihn aber nicht überschreiten (bei der iterativen Anwendung von Abstiegsverfahren), da die Barrieren $-\ln(-g_i(x))$ die Zielfunktion H_δ^L sehr stark vergrößern würden. Der Algorithmus ist also eine Innere-Punkt-Methode. Für konvexe Aufgaben ohne Gleichungsrestriktionen gilt der Satz [20]

Satz 5.2 *Unter den Voraussetzungen*

- (i) *Es seien f und alle g_i konvex und differenzierbar,*
- (ii) *es existiere ein Slater-Punkt, d.h. ein x^0 mit $g_i(x^0) < 0, \forall i$*

gilt: Jeder Häufungspunkt einer konvergenten Teilfolge $\{x(\delta)\}_{\delta \rightarrow +0}$ von Minimalpunkten (für $H_{\delta,I}^L(\cdot)$) löst (2.1).

Die "passende Wahl" der Grössen $\delta_i \rightarrow +0$ bildet ähnliche Anwendungsprobleme wie bei der Strafmethode.

5.2 Mix-Methode

Fall 3. Sei $\varepsilon_i \neq 0, \varepsilon_j \neq 0$, beliebig $\forall i, j$

Analog zu den Fall 1. und Fall 2. kann man verallgemeinern, dass jede Lösung des Kojima-Systems $F^\varepsilon(s) = 0$ einen stationären Punkt für die folgende Funktion liefert (und umgekehrt).

$$H_{\varepsilon,I}^G(x) := f(x) + \frac{1}{2} \sum_{i \in E^+} \frac{1}{\varepsilon_i} g^+(x)^2 + \frac{1}{2} \sum_{i \in E^- \cap I} \frac{1}{\varepsilon_i} g^-(x)^2 + \frac{1}{2} \sum_{j=1}^{\kappa} \frac{1}{\varepsilon_j} h_j(x)^2,$$

wobei $E^+ := \{i | \varepsilon_i > 0\}$ und $E^- := \{i | \varepsilon_i < 0\}$.

5.3 Symmetrische Störung

Offenbar ist es möglich, die Kojima-Funktion (2.2) auf die verschiedenste Art und Weise zu stören. Die Frage ist aber, ob das überhaupt irgendein Sinn hat und wenn ja, wie man dann die Nullstellen interpretieren kann. Wir schauen uns dazu eine Störung an, die in Kapitel 6 eine bedeutende Rolle bei der Untersuchung NCP-Funktionen spielen wird.

Es ist eine symmetrische Störung. Zur Vereinfachung betrachten wir jetzt das Problem (2.1) ohne Gleichungsrestriktionen. Hierzu definieren wir die folgende gestörte Kojima-Gleichung:

$$\begin{aligned} F_1^\varepsilon(x, y) &:= Df(x) + \sum_{i=1}^m (y_i^+ + \varepsilon_i y_i^-) Dg_i(x) = 0 \\ F_{2,i}^\varepsilon(x, y) &:= g_i(x) - y_i^- - \varepsilon_i y_i^+ = 0, \quad i = \overline{1, m}. \end{aligned} \tag{5.10}$$

Wie vorher unterscheiden wir zwei Fälle.

Fall 1. $\varepsilon_i > 0, \forall i$

Sei $F^\varepsilon(x, y) = 0$.

$$\begin{aligned} \text{Wenn } y_i \geq 0, \text{ so folgt } & y_i^- = 0 \quad \text{und} \quad y_i^+ = g_i^+(x)/\varepsilon_i. \\ \text{Wenn } y_i < 0, \text{ so folgt } & y_i^+ = 0 \quad \text{und} \quad y_i^- = g_i^-(x). \end{aligned}$$

Setzen wir y_i in die erste Gleichung aus (5.10) ein, erhalten wir

$$F_1^\varepsilon(s) = Df(x) + \sum_{i=1}^m \left(\frac{1}{\varepsilon_i} g_i^+(x) + \varepsilon_i g_i^-(x) \right) Dg_i(x) = 0.$$

Das bedeutet, dass x ein stationärer Punkt der (Straf-) Funktion

$$H_\varepsilon^S(x) := f(x) + \frac{1}{2} \sum_{i=1}^m \left(\frac{1}{\varepsilon_i} g_i^+(x)^2 + \varepsilon_i g_i^-(x)^2 \right)$$

ist.

Fall 2. $\varepsilon_i < 0, \forall i$

Sei $F^\varepsilon(x, y) = 0$.

$$\begin{aligned} \text{Wenn } y_i \geq 0, \text{ so folgt } & y_i^- = 0 \quad \text{und} \quad y_i^+ = g_i^-(x)/\varepsilon_i. \\ \text{Wenn } y_i < 0, \text{ so folgt } & y_i^+ = 0 \quad \text{und} \quad y_i^- = g_i^-(x). \end{aligned}$$

Die erste Gleichung aus (5.10) sieht nun so aus

$$F_1^\varepsilon(s) = Df(x) + \sum_{i \in I} \frac{1}{\varepsilon_i} g_i^-(x) Dg_i(x) + \sum_{i \notin I} \varepsilon_i g_i^-(x) Dg_i(x) = 0,$$

d.h. x ist auch ein stationärer Punkt der folgenden (Barriere-) Funktion

$$H_{\varepsilon, I}^B(x) := f(x) + \frac{1}{2} \sum_{i \in I} \frac{1}{\varepsilon_i} g_i^-(x)^2 + \frac{1}{2} \sum_{i \notin I} \varepsilon_i g_i^-(x)^2,$$

wobei I nach (5.4) definiert ist.

Analog zu vorher (5.3), (5.6) bekommt man also eine Interpretation für eine solche symmetrische Störung mittels stationärer Punkte von Hilfszielfunktionen. Neu ist nun nur, dass hier in der Funktion $H_\varepsilon^S(x)$ bzw. $H_{\varepsilon, I}^B(x)$ auch die Funktionen g_i mit $i \notin I$ erscheinen.

Elimination der Indexmenge I

Um ein komplettes Bild von den Variationen der Störungen solcher symmetrischen Art zu haben und die Indexmenge I zu eliminieren, führen wir schliesslich eine letzte Funktion ein, die sich von F^ε aus (5.10) nur durch den Betrag in $F_{2,i}^\varepsilon$ unterscheidet:

$$\begin{aligned} F_1^\varepsilon(x, y) &:= Df(x) + \sum_{i=1}^m (y_i^+ + \varepsilon_i y_i^-) Dg_i(x) \\ F_{2,i}^\varepsilon(x, y) &:= g_i(x) - y_i^- - |\varepsilon_i| y_i^+, \quad i = \overline{1, m} \end{aligned}$$

Im Straf–Fall 1. stimmt sie mit (5.10) überein. Also interessiert uns jetzt nur der Barriere–Fall 2, d.h. $\varepsilon_i < 0 \forall i$.

Sei (x, y) eine Lösung von $F^\varepsilon(x, y) = 0$. Man sieht leicht,

$$\begin{aligned} \text{wenn } y_i \geq 0, \text{ so folgt } & y_i^- = 0 \quad \text{und} \quad y_i^+ = g_i^+(x)/|\varepsilon_i|, \\ \text{wenn } y_i < 0, \text{ so folgt } & y_i^+ = 0 \quad \text{und} \quad y_i^- = g_i^-(x). \end{aligned}$$

Daraus folgt, dass die zugeordnete Straf-Barriere-Funktion sich nun ohne Einführung der Indexmenge I schreiben läßt

$$H_\varepsilon^{SB}(x) := f(x) + \frac{1}{2} \sum_{i=1}^m \left(\frac{1}{|\varepsilon_i|} g_i^+(x)^2 - |\varepsilon_i| g_i^-(x)^2 \right)$$

was natürlich die Darstellung vereinfacht.

Nachteilig bleibt aber, dass die Funktion $H_\varepsilon^B(x)$ auch im konvexen Fall in Bezug auf einen *Sattelpunkt* untersucht werden muß.

Allerdings ist mit $g_i(x) = q_i < 0$ der Punkt x wegen

$$\frac{1}{2} D(g_i(x)^-)^2 = g_i(x)^- Dg_i(x) = g_i(x)^2 D(\ln(-g_i(x))) = q_i^2 D(\ln(-g_i(x)))$$

auch ein stationärer Punkt der Funktion

$$H_\varepsilon^{SB}(x) := f(x) + \frac{1}{2} \sum_{i:g_i(x)>0} \frac{1}{|\varepsilon_i|} g_i(x)^2 - \frac{1}{2} \sum_{i:g_i(x)<0} |\varepsilon_i| q_i^2 \ln(-g_i(x)).$$

Hier realisiert er für konvexe Probleme wieder ein *Minimum*.

5.4 Abschätzungen

Aus der Form aller hier angegebenen (gestörten) Kojima-Funktionen ist leicht zu sehen, dass

$$\|F^\varepsilon - F\|_{B(s^*, \rho)}^{0,1} \leq K \|\varepsilon\|, \quad (5.11)$$

wobei die Konstante K nur von ρ und $\|(y^*, z^*)\|$ abhängt.

Das erlaubt uns, als wesentliche Motivation dafür, solche Störungen überhaupt zu betrachten, die Abschätzung der entsprechenden Nullstellen (also auch der stationären Punkte der Straf- und Barrierefunktionen) auf der Grundlage bekannten Regularitätseigenschaften in Kapitel 3 und nichtlinearer Störungen der jeweiligen Gleichung.

So folgt nun (für Nullstellen $s(\varepsilon)$ von F^ε in einer gewissen Kugel $B(s^*, \rho)$ und normkleine Parameter $\varepsilon, \varepsilon'$) unter der strengen Regularität von F in s^* sowohl die Existenz von $s(\varepsilon) \in B(s^*, \rho)$ als auch

$$\|s(\varepsilon) - s(\varepsilon')\| \leq C \|F^\varepsilon - F^{\varepsilon'}\|_{B(s^*, \rho)}^{0,1} \leq CK \|\varepsilon - \varepsilon'\|, \quad (5.12)$$

wobei C eine Lipschitz-Konstante ist.

Außerdem sind die Lösungen $s(\varepsilon), s(\varepsilon') \in B(s^*, \rho)$ eindeutig. Auf diese Weise kann man einerseits direkt die Konvergenz der oben genannten Methoden abschätzen (mit $\varepsilon' = 0$ in (5.12)). Andererseits kann man auch zwei beliebige dieser Verfahren unterschiedlichen Typs miteinander vergleichen (z.B. Barriere- und Strafmethode), weil in (5.12) weder das Vorzeichen der $\varepsilon_i, \varepsilon'_i$ noch die konkrete Form von F^ε eine Rolle spielt. Man benötigt für die Abschätzung nur (5.11). Ganz analog lassen sich die anderen Regularitätseigenschaften von F ausnutzen. Auch ein für konvexe Optimierungsaufgaben sinnvoller "Regularisierungsterm" der Form εx , $\varepsilon > 0$ in der Lagrange-Gleichung läßt noch dieselben Abschätzungen zu und entspricht einer Tychonov-Regularisierung im Sinne von [11].

Konvergenz von δ_i

Wie es schon erwähnt wurde (5.8), fordert die Barrieremethode dass $|\delta_i| \rightarrow 0$, wenn $|\varepsilon_i| \rightarrow 0$. Um dies zu sehen für $\delta_i := -g_i(x)^2/\varepsilon_i$ (nach (5.7)), benutzt man zuerst, dass die Funktion $g(x)$ Lipschitz-stetig nach Voraussetzung ist. Die strenge Regularität sagt uns dann, dass $x(\cdot)$ ebenfalls Lipschitz-stetig ist. Wir erhalten also

$$|g_i(x(\varepsilon)) - g_i(x^*)| \leq L|\varepsilon|.$$

Für die aktiven Indizes $i \in I$, die uns eigentlich interessieren, gilt allgemein $g_i(x(\varepsilon)) = \varepsilon_i y_i^+$ und mit $\varepsilon \rightarrow -0$ folgt $g_i(x^*) = 0$. Die vorige Lipschitz-Abschätzung sieht dann folgendermaßen aus:

$$|g_i(x(\varepsilon))| \leq L|\varepsilon|, \quad \forall i \in I$$

Daraus folgt (siehe (5.7)): $|\delta_i| \leq L^2|\varepsilon|^2/|\varepsilon_i|$. Somit ist $|\delta_i| \rightarrow 0$ gesichert, wenn nur ε so festgelegt wird, dass $|\varepsilon|$ wenigstens in derselben Ordnung wie alle $|\varepsilon_i|$ gegen 0 konvergieren (d.h. ε darf nicht asymptotisch entlang einer Koordinatenachse gegen den Ursprung konvergieren).

5.5 Anwendung gestörter Kojima-Funktionen im Newton-Verfahren

Sei s^* eine streng reguläre Nullstelle von (2.2). Um die Nachteile der direkten Anwendung des Newton-Verfahrens auf die Kojima-Funktion zu umgehen, wähle man $\varepsilon = \varepsilon(t)$ in Abhängigkeit vom aktuellen Iterationspunkt s^t , $t = 0, 1, 2, \dots$, wobei der Startpunkt s^0 nahe s^* sei. Man kann dann

1) Newton-Schritte direkt auf $F^{\varepsilon(t)}$ anwenden

oder auch

2) gleichzeitig mit der Newton-Matrix von $F^{\varepsilon(t)}$ und dem exakten Wert der ungestörten Kojima-Funktion F arbeiten.

Den Abschätzungen für das gestörte Newton-Verfahren folgend, kann man ε im zweiten Fall grösser wählen, was bei hohen Genauigkeiten, d.h. bei normkleinem $F(s^t)$, die Gefahr von Auslöschungseffekten verringert. Mit ε_j bezeichnen wir jeweils die Störungen, die den Gleichungsrestriktionen $h_j = 0$ entsprechen.

Methode 1

Wir setzen:	zum Beispiel:
$\varepsilon_i > 0$, wenn $g_i > 0$	$\varepsilon_i(t) := \text{sign}(g_i(x))\ F(s^t)\ ^2$
$\varepsilon_i < 0$, wenn $g_i < 0$	
$\varepsilon_j > 0$	$\varepsilon_j(t) := \ F(s^t)\ ^2$

Danach wendet man, mit entsprechendem $\varepsilon(t)$, einen Newton-Schritt auf das jeweilige gestörte Kojima-System an, um s^{t+1} zu bestimmen. Die strenge Regularität (3.4) von F sichert dann mittels der obigen Abschätzungen sowohl die Existenz und Eindeutigkeit einer Nullstelle von $F^{\varepsilon(t)}(s) = 0$ als auch (über Newton-Verfahren unter Störungen) die Konvergenz $s^{t+1} \rightarrow s^*$.

Methode 2

Wir setzen:	zum Beispiel:
$\varepsilon_i > 0$, wenn $g_i > 0$	$\varepsilon_i(t) := \text{sign}(g_i(x))\ F(s^t)\ $
$\varepsilon_i < 0$, wenn $g_i < 0$	
$\varepsilon_j > 0$	$\varepsilon_j(t) := \ F(s^t)\ $

Danach wendet man, mit entsprechendem $\varepsilon(t)$, einen Newton–Schritt der Form

$$F(s^t) + TF^{\varepsilon(t)}(s^t)(\sigma) = 0$$

an, um $s^{t+1} = s^t + \sigma$ zu setzen.

Lösbarkeit und Konvergenz sind erneut durch das Newton–Verfahren unter Störungen gesichert.

Die einzelnen Newton–Schritte sind nun als Straf- bzw. Barriere–Schritte interpretierbar und können so — bei schlechten Startpunkten — auch durch mehrere Schritte erster Ordnung angewandt auf die jeweilige Hilfszielfunktion ersetzt werden (um anschliessend wieder zum Newton–Verfahren überzugehen).

Kapitel 6

Einheitliche Darstellung von NCP–Ansätzen

In diesem Kapitel wollen wir uns mit NCP[†]–Funktionen für die Lösung von Komplementaritätsproblem ausführlicher beschäftigen. Das NCP–Problem

$$u(x) \geq 0, \quad v(x) \geq 0, \quad \langle u, v \rangle = 0 \quad (6.1)$$

haben wir schon eingeführt (2.14) und haben gezeigt, wie man es mittels verallgemeinerten Kojima–Funktion darstellen kann. Wir werden hier mehr auf die praktische Seite für das Lösen der Aufgabe eingehen. Der Schwerpunkt unserer Betrachtung ist die NCP–Funktion Methode und ihr Zusammenhang mit dem Kojima–System. Dabei benutzen wir in den beiden ersten Abschnitten die Grundlagen aus [14].

6.1 NCP–Funktion

Definition 6.1 Eine Funktion $G(s, t) : \mathbb{R}^2 \rightarrow \mathbb{R}$ heißt NCP–Funktion, falls

$$G^{-1}(0) = \{(s, t) \mid s \geq 0, t \geq 0, st = 0\}.$$

Die Definition ermöglicht uns die äquivalente Beschreibung des Originalproblems (6.1) mittels der Gleichung

$$h(x) = 0; \quad h_i(x) = G(z_i(x)) \quad \forall i = 1, \dots, n \quad (6.2)$$

wobei

$$z_i(x) = (u_i(x), v_i(x)), \quad i = 1, \dots, n.$$

Es gibt verschiedene Möglichkeiten, die Aufgabe (6.2) zu lösen. In der Praxis verwendet man meistens die folgende Zugänge:

- (i) eine sogenannte Merit–Funktion wird minimiert, z.B.

$$\frac{1}{2} \sum_{i=1}^n h_i^2(x)$$

- (ii) man versucht (6.2) mittels einer Newton–Methode direkt zu lösen

[†]Abkürzung von 'Nonlinear Complementarity Problem'

(iii) Kombination von (i) und (ii)

In jedem dieser Fälle sollte G gewisse Glattheitseigenschaften besitzen und möglichst einfach sein. Mit Blick auf das Newton-Verfahren ist es allerdings nicht günstig, wenn G vom Typ C^1 nahe $(0,0)$ ist. Denn dann würden die partiellen Ableitungen $G_s(0,0)$ und $G_t(0,0)$ wegen der Struktur von G beide verschwinden (auf den positiven Halbachsen ist G konstant, also jeweils eine partielle Ableitung Null), und es würde folgen

$$DG(z_i(x)) = 0 \quad \text{und} \quad Dh_i(x) = DG(z_i(x))Dz_i(x) = 0,$$

falls $z_i(x) = (0,0)$ und $z \in C^1$. Damit wäre $Dh(x)$ nicht regulär und das Newton-Verfahren (auch mit verallgemeinerten Ableitungen) im Punkt x nicht anwendbar. Also sollte G im Punkt $(0,0)$ nicht differenzierbar sein. Deswegen setzen wir voraus, dass G positiv homogen ist. Solche Funktionen haben nämlich die benötigte Eigenschaft.

Andererseits, damit das Newton-Verfahren überhaupt angewandt werden kann, braucht man eine gewisse Glattheit und Monotonie von G . Wir fordern: G sei pseudo-glatt und besitze auf der Einheitskugel höchstens endlich viele Punkte p^k , die nicht zur C^1 -Menge $\Theta^1(G)$ gehören.

Als pseudo-glatte Funktion ist G dann lokal Lipschitz und gehört weiter zur Klasse $LocPC^1$. Die Polyeder P^s in der Definition sind hier Kegel der Form $K \cap O$, wobei O einen der vier Orthanten bezeichnet und die Kegel K irgendwelche spitze Kegel sind mit Scheitel im Ursprung, so dass deren Kantenvektoren insgesamt alle $p^k \in \mathbb{R}^n \setminus \Theta^1(G)$ enthalten.

Die positive Homogenität (6.6) liefert in jedem C^1 -Punkt von G

$$DG(\sigma) = DG(\lambda\sigma), \quad \forall \lambda > 0, \quad \sigma \in \Theta^1(G), \quad (6.3)$$

und sagt uns, dass solche NCP-Funktionen für $(s,t) \in \Theta^1(G)$ eine besondere Form haben:

$$G(s,t) = DG(s,t)(s,t)^T = sG_s(s,t) + tG_t(s,t). \quad (6.4)$$

Außerdem folgt

$$D^0G(0) = \text{cl } DG(\Theta^1(G)), \quad (6.5)$$

denn es gilt für $\sigma_k \in \Theta^1(G), \forall k$ immer:

$$\lim_{\sigma_k \rightarrow \sigma^*} DG(\sigma_k) = \lim_{\lambda_k \sigma_k \rightarrow \lambda^* \sigma^*} DG(\lambda_k \sigma_k) = \lim_{\lambda_k \sigma_k \rightarrow 0} DG(\lambda_k \sigma_k) = \lim_{\sigma_k \rightarrow 0} DG(\sigma_k),$$

wobei $\lim_{k \rightarrow \infty} \lambda_k = \lambda^*$, $\lambda_k \neq 0, \forall k$ und $\lambda^* = 0$.

Wir setzen weiter voraus, dass die positiven Halbachsen (sie enthalten $z_i(x^*) = (u_i(x^*), v_i(x^*))$) für streng komplementäre Lösungen x^* in $\Theta^1(G)$ liegen und dort die Ableitung DG nicht verschwindet. Betrachtet man die positiv-homogene Funktion G auf der Kugel, sieht man dann sofort, dass G im Inneren des positiven Orthanten entweder positiv und dann im Inneren der übrigen Orthanten negativ ist oder umgekehrt. Außerdem sei G monoton, d.h.

$$0 \neq DG(s,t) \geq 0, \quad \forall (s,t) \in \Theta^1(G).$$

Zusammenfassend verlangen wir also:

$$G \text{ ist positiv homogen, d.h. } G(\lambda\sigma) = \lambda G(\sigma) \quad \forall \sigma \in \mathbb{R}^2, \lambda \geq 0, \quad (6.6)$$

die C^1 –Menge von G , $\Theta^1(G)$ enthält $G^{-1}(0) \setminus \{0\}$ (6.7)

$$DG(\sigma) \geq 0 \quad \text{und} \quad DG(\sigma) \neq 0, \quad \forall \sigma \in \Theta^1(G) \quad (6.8)$$

sowie

$$G \in \text{LocPC}^1. \quad (6.9)$$

Die Klasse dieser Funktionen bezeichnen wir wie in [14] mit $pNCP$. In der Arbeit wird sehr oft auch die folgende Eigenschaft benutzt.

Definition 6.2 *Wir nennen eine Funktion $G \in pNCP$ streng monoton, wenn*

$$a > b \quad \forall (a, b) \in D^0G(\sigma) \quad \text{und} \quad \sigma \in \mathbb{R}^2 \setminus G^{-1}(0).$$

Tatsächlich umfasst $pNCP$ nicht alle in der Literatur studierten NCP Funktionen, siehe etwa [30]; z.B. werden mitunter noch Terme höherer Ordnung wie $(s^+t^+)^2$, addiert. Sie sind heuristisch motiviert für globale Konvergenzeigenschaften (Verstärkung von Monotonie), haben jedoch keinen qualitativen Einfluss auf das Konvergenzverhalten von Newton–Verfahren nahe einer Lösung, da die (verallgemeinerten) Ableitungen sich nur in der Ordnung $O(x - x^*)$ unterscheiden und die Funktionswerte selbst von der Ordnung $o(x - x^*)$ (siehe Satz 4.3).

Mit der oft benutzten $pNCP$ –Funktion (siehe [5, 9, 10])

$$G(s, t) = s + t - \sqrt{s^2 + t^2},$$

erhält man z.B. für $(s, t) \neq 0$

$$G_s = 1 - \frac{s}{\sqrt{s^2 + t^2}}, \quad G_t = 1 - \frac{t}{\sqrt{s^2 + t^2}}.$$

Mit der Funktion $G(s, t) = \min\{s, t\}$ gilt einfach

$$DG(s, t) = (1, 0)(s < t), \quad DG(s, t) = (0, 1)(s > t).$$

Die erste ist streng monoton, die zweite nicht.

Wichtig für Newton–Verfahren ist, dass die Ableitungen von G auch im Limes nicht verschwinden.

Lemma 6.1 *Für jede Funktion $G \in pNCP$ existiert ein $p > 0$, so dass $\|DG(\sigma)\| \geq p$, $\forall \sigma \in \Theta^1(G)$ gilt, und damit auch $\inf \|D^0G(0)\| \geq p$. Darüberhinaus hat der Gradient von G die folgende Eigenschaft: $DG(e^1) = \lambda e^2$ und $DG(e^2) = \mu e^1$, wobei e^1, e^2 die Einheitsvektoren im Raum \mathbb{R}^2 , und $\lambda, \mu > 0$ sind.*

Beweis:

Angenommen

$$\lim DG(\sigma_k) = 0 \quad (6.10)$$

gilt für eine Folge $\sigma_k \in \Theta^1(G)$. Dann ist $\sigma_k \neq 0$, und wegen Norminvarianz der Ableitung (6.3) kann $\|\sigma_k\| = 1 \quad \forall k$ angenommen werden. Somit konvergiert eine unendliche Teilfolge (o.B.d.A.

schon die gesamte Folge) gegen ein Element $\sigma^* \in \text{bd } B$. Ist $\sigma^* \in \Theta^1(G)$, so folgt mit (6.8) der Widerspruch

$$0 = \lim DG(\sigma_k) = DG(\sigma^*) \neq 0.$$

Sei also $\sigma^* \notin \Theta^1(G)$. Dann kann σ^* wegen (6.7) auch nicht auf den beiden positiven Halbachsen liegen. Deshalb ist $G(\sigma^*) \neq 0$, und wir erhalten aus (6.10) und (6.4) den Widerspruch

$$0 \neq G(\sigma^*) = \lim_k G(\sigma_k) = \lim_k DG(\sigma_k)\sigma_k = 0.$$

Dank (6.4) haben wir auch

$$G(e^1) = DG(e^1)^T e^1 = 0$$

Aus (6.7) und (6.8) folgt so wegen $G_s(e^1) = 0$ weiter die Behauptung $DG(e^1) = \lambda e^2$ mit $\lambda > 0$. Für e^2 ist alles analog. ■

6.2 Transformation der Iterierten beim Newton-Verfahren

Der Newton-Schritt bei Anwendung einer $pNCP$ -Funktion

Wir können nun einen Newton-Schritt für die Aufgabe (6.2) einfach mittels der partiellen Ableitungen von G interpretieren. Zu lösen ist

$$h_i(x) + Dh_i(x)\xi = 0 \quad i = 1, \dots, n \quad ; \quad x_{new} = x + \xi.$$

Mögen die entsprechenden Frechét-Ableitungen von G existieren und sei $z = (u, v) \in C^1$. Dann erhalten wir also

$$G(z_i(x)) + G_s(z_i(x))Du_i(x)\xi + G_t(z_i(x))Dv_i(x)\xi = 0$$

oder mit (6.4)

$$G_s(z_i(x))[u_i(x) + Du_i(x)\xi] + G_t(z_i(x))[v_i(x) + Dv_i(x)\xi] = 0.$$

Setzt man

$$a_i := G_s(z_i(x)) \quad \text{und} \quad b_i := G_t(z_i(x)), \quad (6.11)$$

so bedeutet die letzte Gleichung, dass $a_i \geq 0$ und $b_i \geq 0$ Bewichtungen der beiden Linearisierungen von u_i bzw. v_i im Punkt x sind:

$$a_i[u_i(x) + Du_i(x)\xi] + b_i[v_i(x) + Dv_i(x)\xi] = 0.$$

Mit $Lu_i(x, \xi) = u_i(x) + Du_i(x)\xi$ und wegen $(a_i + b_i) \neq 0$ läßt sich diese Gleichung schließlich so umschreiben

$$(1 - k_i)Lu_i(x, \xi) + k_iLv_i(x, \xi) = 0, \quad (6.12)$$

wobei

$$k_i = \frac{b_i}{a_i + b_i} = \frac{G_t(z_i(x))}{G_s(z_i(x)) + G_t(z_i(x))}. \quad (6.13)$$

Aus (6.8) folgt sofort $k_i \in [0, 1]$. Die Gleichung (6.12) ist dann die Newton-Gleichung (1.7) für unser konkretes Komplementaritätsproblem (6.1).

Nach (6.3) und Lemma 6.1 gilt hierbei mit einem gewissen $\rho > 0$

$$\begin{aligned} a_i = 0, b_i \geq \rho, k_i = 1 & \text{ wenn } u_i(x) > 0, v_i(x) = 0 \text{ und} \\ b_i = 0, a_i \geq \rho, k_i = 0 & \text{ wenn } v_i(x) > 0, u_i(x) = 0. \end{aligned} \quad (6.14)$$

Bemerkung 6.1 Die Überlegungen (6.12), (6.14), und damit $(a_i + b_i) \neq 0$, bleiben (wegen positiver Homogenität und nach Grenzübergang $\sigma_i \in \Theta^1(G)$, $\sigma_i \rightarrow z_i(x)$) auch dann bestehen, wenn man nur voraussetzt, dass

$$(a_i, b_i) \in D^0G(z_i(x))$$

gilt. Ebenso kann man $Dz(x)$ durch Elemente $\zeta \in D^0z(x)$ ersetzen, wenn nur $z \in \text{loc}PC^1$ gilt.

Transformation der Iterierten

Wie es schon früher erwähnt wurde (siehe Abschnitt 2.3.1), kann man das Komplementaritätsproblem in die Form eines Kojima–Systems transformieren:

$$\begin{aligned} F_i(x, y) &:= u_i(x) - y_i^+ = 0 \\ F_{n+i}(x, y) &:= -v_i(x) - y_i^- = 0 \\ i &= 1, \dots, n. \end{aligned} \tag{6.15}$$

Um die Lösung von $F(x, y) = 0$ zu finden, wollen wir hier das Newton–Verfahren anwenden. Die entsprechenden Newton–Gleichungen sehen dann konkret (bei Benutzung der Ableitung TF) so aus

$$\begin{aligned} F_i(x, y) + Du_i(x)\xi - r_i\eta_i &= 0 \\ F_{n+i}(x, y) - Dv_i(x)\xi - (1 - r_i)\eta_i &= 0 \\ i &= 1, \dots, n \end{aligned} \tag{6.16}$$

wobei

$$r \in R_T(y) \tag{6.17}$$

Die Menge $R_T(y)$ ist durch (3.6) definiert ($m = n$).

Lemma 6.2 [14]

(i) Ist (ξ, η) eine Lösung der Newton–Gleichung (6.16) für den Iterationspunkt (x, y) mit $r \in R_T(y)$, so gilt (6.12) mit $k_i = r_i$, $\forall i$.

(ii) Umgekehrt, wenn ξ die (NCP–Newton–) Gleichungen (6.12) im Iterationspunkt x für ein $k = r \in R_T(y)$, $y = u(x) - v(x)$ erfüllt, so ist (ξ, η) mit

$$\eta_i = \begin{cases} F_i(x, y) + Du_i(x)\xi, & \text{falls } r_i = 1 \\ F_{n+i}(x, y) - Dv_i(x)\xi, & \text{falls } r_i = 0 \\ (F_i(x, y) + Du_i(x)\xi)/r_i, & \text{sonst} \end{cases}$$

eine Lösung von (6.16).

Beweis:

(i) Zur Abkürzung sei $F := F(x, y)$. Aus (6.16) haben wir

$$F_i + Du_i(x)\xi = r_i\eta_i, \quad F_{n+i} - Dv_i(x)\xi = (1 - r_i)\eta_i$$

Weiter folgt

$$(1 - r_i)F_i + (1 - r_i)Du_i(x)\xi = (1 - r_i)r_i\eta_i = r_iF_{n+i} - r_iDv_i(x)\xi$$

oder

$$(1 - r_i)F_i + (1 - r_i)Du_i(x)\xi - r_iF_{n+i} + r_iDv_i(x)\xi = 0$$

Setzen wir F_i und F_{n+i} ein und nutzen $r_i y_i^- = (1 - r_i) y_i^+$ (da $r \in R_T(y)$) aus, so erhalten wir die Gleichung

$$(1 - r_i)[u_i(x) + Du_i(x)\xi] + r_i[v_i(x) + Dv_i(x)\xi] = 0$$

Folglich gilt (6.12), sobald wir $k_i = r_i$ einsetzen.

(ii) Wir haben drei Fälle zu betrachten:

- 1) $r_i = 1$ dann $y_i^- = 0$, $r_i \eta_i = \eta_i = F_i + Du_i(x)\xi$
und aus (6.12) folgt $v_i(x) + Dv_i(x)\xi = 0$,
d.h. $F_{n+i} - Dv_i(x)\xi =$
 $= F_{n+i} + v_i(x) = -v_i(x) - y_i^- + v_i(x) = 0 = (1 - r_i)\eta_i$
- 2) $r_i = 0$ dann $y_i^+ = 0$, $(1 - r_i)\eta_i = \eta_i = F_{n+i} - Dv_i(x)\xi$
und aus (6.12) folgt $u_i(x) + Du_i(x)\xi = 0$, d.h.
 $F_i + Du_i(x)\xi = F_i - u_i(x) = u_i(x) - y_i^+ - u_i(x) = 0 = r_i \eta_i$
- 3) $0 < r_i < 1$ dann $y_i = 0$, $r_i \eta_i = F_i + Du_i(x)\xi$ und
 $F_{n+i} - Dv_i(x)\xi = -v_i(x) - Dv_i(x)\xi =$
 $= -Lv_i(x, \xi) = r_i^{-1}(1 - r_i)Lu_i(x, \xi) =$
 $= r_i^{-1}(1 - r_i)(F_i + Du_i(x)\xi) = (1 - r_i)\eta_i$

Also ist (6.16) in jedem Fall erfüllt. ■

Das Lemma zeigt, dass sich der enge Zusammenhang zwischen Kojima-Funktion und NCP-Funktionen auch auf das Newton-Verfahren erstreckt. Der einzige Unterschied zwischen beiden (Newton-) Systemen (6.12) und (6.16) liegt in der Definition der Konstanten k_i und r_i . Für k_i gilt (6.13), womit $0 < k_i < 1$ auch mit $y_i := u_i(x) - v_i(x) \neq 0$ gelten darf. Für r_i dagegen soll gelten

$$r \in R_T(y) \text{ mit } y = u(x) - v(x),$$

was $0 < r_i < 1$ nur für $y_i := u_i(x) - v_i(x) = 0$ zulässt.

Wird k_i mittels D^0G über (6.13) definiert, so folgt $k \in R_T(y)$ und $k_i \in \{0, 1\}$ mit $y = u(x) - v(x)$, sofern man die einfachste NCP-Funktion $G(s, t) = \min\{s, t\}$ benutzt. Mit einer beliebigen p NCP-Funktion G wird $k \in R_T(y)$ mit $y = u(x) - v(x)$ zumindest in jeder Lösung $x = x^*$ des Komplementaritätsproblems erfüllt (wegen (6.14)).

Weiter gilt mit $k = k(x)$ nach (6.13) im Falle der Konvergenz $x \rightarrow x^*$, $k(x) \rightarrow k^*$ im Limes auch $k^* \in R_T(y)$. Allerdings ist allgemein, d.h., für beliebige p NCP-Funktionen, die Inklusion $k(x) \in R_T(u(x) - v(x))$ nicht gesichert, wenn $k(x)$ nach (6.13) gebildet wird.

Die durch G definierten Newton-Schritte können aber, wie wir gleich sehen, durch kleine Änderungen von F modelliert werden. Anderes gesagt, den Unterschied zwischen r und k wird beseitigt.

6.3 Symmetrische Störung des Kojima-Systems und p NCP-Funktionen im Newton-Verfahren

Wir betrachten an Stelle des Systems (6.15) ein folgendermaßen symmetrisch (durch die Terme $\varepsilon_i y_i^-$ bzw. $\varepsilon_i y_i^+$) gestörtes Kojima-System:

$$\begin{aligned} F_i^\varepsilon(x, y) &:= u_i(x) - y_i^+ - \varepsilon_i y_i^- = 0 \\ F_{n+i}^\varepsilon(x, y) &:= -v_i(x) - y_i^- - \varepsilon_i y_i^+ = 0 \\ i &= 1, \dots, n \end{aligned} \tag{6.18}$$

wobei $\varepsilon \in \mathbb{R}^n$ ein beliebiger Parameter ist. Wenden wir auf $F^\varepsilon(x, y)$ das (TF^-) [†] Newton–Verfahren an, erhalten wir die Gleichungen

$$\begin{aligned} F_i^\varepsilon(x, y) + Du_i(x)\xi - (\varepsilon_i + r_i - \varepsilon_i r_i)\eta_i &= 0 \\ F_{n+i}^\varepsilon(x, y) - Dv_i(x)\xi - (1 - r_i + \varepsilon_i r_i)\eta_i &= 0 \\ i &= 1, \dots, n, \end{aligned} \tag{6.19}$$

die für ein $r \in R_T(y)$ zu erfüllen sind.

Im folgenden Satz untersuchen wir den allgemeinen Zusammenhang zwischen den Systemen (6.12) und (6.19). Wir werden sehen, wie dieser Zusammenhang sich wesentlich vereinfacht, wenn wir nur Elemente $r \in R_T(y)$ mit $r_i \in \{0, 1\}$, $\forall i$ zulassen und wenn die $pNCP$ –Funktion G streng monoton ist.

6.3.1 Beliebige $pNCP$ –Funktionen

Im folgenden Satz spielt die nachstehende Definition von ε mit $r \in R_T(y)$ die entscheidende Rolle

$$\varepsilon_i = \frac{(a_i + b_i)r_i - b_i}{(a_i + b_i)r_i - a_i}, \quad \forall i. \tag{6.20}$$

Man sieht gleich, dass $\varepsilon_i \geq 0, \forall i$ gilt. Wenn wir mittels des symmetrisch gestörten Kojima–Systems (6.18) zu einer Lösung des original NCP (6.1) kommen wollen, müssen wir nach wie vor den Parameter ε gegen Null konvergieren lassen.

Die angegebene Definition von ε_i (6.20) bewirkt für die Konvergenz $x \rightarrow x^*$ gegen eine Lösung, dass $\varepsilon_i \rightarrow 0$ falls in Komponente i strikte Komplementarität $u_i(x^*) + v_i(x^*) > 0$ vorliegt. Tatsächlich:

Situation I (strikte Komplementarität): Sei $(u_i(x^*), v_i(x^*)) \neq (0, 0)$.

Fall 1. $v_i(x^*) > 0$.

D.h. $u_i(x) \rightarrow 0$ und nach (6.14) folgt $\varepsilon_i = b_i(x)/a_i(x) \rightarrow 0$.

Fall 2. $u_i(x^*) > 0$.

D.h. $v_i(x) \rightarrow 0$ und nach (6.14) folgt $\varepsilon_i = a_i(x)/b_i(x) \rightarrow 0$.

Situation II: Sei $(u_i(x^*), v_i(x^*)) = (0, 0)$. Dann kann ε_i einen positiven endlichen Grenzwert besitzen; z.B. mit $G = s + t - \sqrt{s^2 + t^2}$ und $u_i(x) = v_i(x) \rightarrow 0$. Die Beschränktheit erkennt man durch Betrachtung von DG bzw. D^0G auf der (kompakten) Sphäre. Weil nun $y_i^* = u_i(x^*) - v_i(x^*) = 0$ ist, bleibt allerdings (x^*, y^*) auch mit jedem ε_i eine Lösung von System (6.18).

Wir haben also gezeigt, dass die Wahl von ε nach (6.20) legitim ist und die Regularitätsaussagen damit weiterhin gelten.

Satz 6.1 Für festes $x \in \mathbb{R}^n$ seien a_i, b_i nach (6.11) erklärt. Wir setzen

$$(a_i + b_i)r_i - a_i \neq 0 \tag{6.21}$$

voraus, und betrachten ε aus (6.20). Dann gelten die folgenden Aussagen:

(i) Ist (ξ, η) eine Lösung von (6.19) im Punkt (x, y) , mit $r \in R_T(y)$ und gilt außerdem $r_i \neq \frac{1}{2}$, so

[†]Mit Thibault–Ableitung

erfüllen ξ und x die (NCP-) Gleichungen (6.12), (6.13).

(ii) Erfüllt ξ und x die Gleichungen (6.12) mit $b_i \neq 0$ und definiert man η mit Hilfe von $y = u(x) - v(x)$ und $r \in R_T(y)$, $r_i \neq \frac{1}{2}$ als

$$\eta_i = \begin{cases} F_{n+i}^\varepsilon(x, y) - Dv_i(x)\xi, & \text{falls } r_i = 0 \\ (F_i^\varepsilon(x, y) + Du_i(x)\xi)/(\varepsilon_i + r_i - \varepsilon_i r_i), & \text{sonst,} \end{cases}$$

so ist (ξ, η) eine Lösung des symmetrisch gestörten (verallgemeinerten) Kojima-Newton-Systems (6.19) mit ε nach (6.20).

Beweis:

(i) Aus (6.19) haben wir

$$F_i^\varepsilon + Du_i\xi = (\varepsilon_i + r_i - \varepsilon_i r_i)\eta_i, \quad F_{n+i}^\varepsilon - Dv_i\xi = (1 - r_i + \varepsilon_i r_i)\eta_i.$$

Wir multiplizieren die erste Gleichung mit $(1 - r_i + \varepsilon_i r_i)$ und die zweite mit $(\varepsilon_i + r_i - \varepsilon_i r_i)$. Da die linke Seiten jetzt gleich sind, folgt also

$$(1 - r_i + \varepsilon_i r_i)F_i^\varepsilon + (1 - r_i + \varepsilon_i r_i)Du_i\xi - (\varepsilon_i + r_i - \varepsilon_i r_i)F_{n+i}^\varepsilon + (\varepsilon_i + r_i - \varepsilon_i r_i)Dv_i\xi = 0.$$

Setzen wir F_i^ε und F_{n+i}^ε ein und nutzen (da $r \in R_T(y)$)

$$(\varepsilon_i + r_i - \varepsilon_i r_i)(y_i^- + \varepsilon_i y_i^+) = (1 - r_i + \varepsilon_i r_i)(y_i^+ + \varepsilon_i y_i^-)$$

aus, so erhalten wir die Gleichung

$$(1 - r_i + \varepsilon_i r_i)[u_i(x) + Du_i\xi] + (\varepsilon_i + r_i - \varepsilon_i r_i)[v_i(x) + Dv_i\xi] = 0. \quad (6.22)$$

Setzt man nun (6.20) ein, folgt

$$1 - r_i + \varepsilon_i r_i = \frac{a_i(2r_i - 1)}{(a_i + b_i)r_i - a_i}$$

und

$$\varepsilon_i + r_i - \varepsilon_i r_i = \frac{b_i(2r_i - 1)}{(a_i + b_i)r_i - a_i} \quad (6.23)$$

Da $r_i \neq \frac{1}{2}$ und (6.21), $\forall i$ vorausgesetzt ist, schreibt sich (6.22) so um

$$a_i Lu_i(x, \xi) + b_i Lv_i(x, \xi) = 0$$

D.h. die Gleichung (6.12) mit (6.13) ist erfüllt.

(ii) Wegen $r_i \neq \frac{1}{2}$, $b_i \neq 0$ und (6.23) ist auch $\varepsilon_i + r_i - \varepsilon_i r_i \neq 0$. Folglich ist η_i korrekt definiert. Sei also $r \in R_T(y)$, wobei $y = u(x) - v(x)$. Wir haben die folgenden drei Fälle zu betrachten:

- 1) $r_i = 1$ dann sind $y_i^- = 0, \varepsilon_i = a_i/b_i$ und
 $(\varepsilon_i + r_i - \varepsilon_i r_i)\eta_i = F_i^\varepsilon + Du_i\xi$.
 Aus (6.12) folgt $Dv_i\xi = -v_i - a_i/b_i Lu_i(x, \xi)$, d.h.
 $F_{n+i}^\varepsilon - Dv_i\xi = F_{n+i}^\varepsilon + v_i + a_i/b_i Lu_i(x, \xi) = -v_i - y_i^- -$
 $-\varepsilon_i y_i^+ + v_i + a_i/b_i (u_i + Du_i\xi) = -\varepsilon_i y_i^+ + a_i/b_i (\eta_i + y_i^+) =$
 $= \varepsilon_i \eta_i = (1 - r_i + \varepsilon_i r_i)\eta_i$.
- 2) $r_i = 0$ dann sind $y_i^+ = 0, \varepsilon_i = b_i/a_i$ und
 $(1 - r_i + \varepsilon_i r_i)\eta_i = \eta_i = F_{n+i}^\varepsilon - Dv_i\xi$.
 Aus (6.12) folgt $Du_i\xi = -u_i - b_i/a_i Lv_i(x, \xi)$, d.h.
 $F_i^\varepsilon + Du_i\xi = F_i^\varepsilon - u_i - b_i/a_i Lv_i(x, \xi) = u_i(x) - y_i^+ - \varepsilon_i y_i^-$
 $-u_i(x) - b_i/a_i (v_i + Dv_i\xi) = -\varepsilon_i y_i^- + b_i/a_i (\eta_i + y_i^-) =$
 $= \varepsilon_i \eta_i = (\varepsilon_i + r_i - \varepsilon_i r_i)\eta_i$.
- 3) $0 < r_i < 1$ dann sind $y_i = 0, (\varepsilon_i + r_i - \varepsilon_i r_i)\eta_i = F_i^\varepsilon + Du_i\xi$.
 Weiter $F_{n+i}^\varepsilon - Dv_i\xi = -v_i - Dv_i\xi = -Lv_i(x, \xi) =$
 $= a_i/b_i Lu_i(x, \xi) = a_i/b_i (F_i^\varepsilon + Du_i\xi) =$
 $= a_i/b_i (\varepsilon_i + r_i - \varepsilon_i r_i)\eta_i = (1 - r_i + \varepsilon_i r_i)\eta_i$.

Also ist (6.19) in jedem Fall erfüllt. ■

In den weiteren Bemerkungen betrachten wir die Voraussetzungen des Satzes genauer.

Bemerkung 6.2 Die Beschränkung auf $r_i \neq \frac{1}{2}$ kann einfach beseitigt werden, indem wir allgemein festlegen

$$r_i = 1, \text{ wenn } y_i = 0 \quad (6.24)$$

Da wir freie Wahl $r_i \in [0, 1]$ haben, falls $y_i = 0$, ist das legitim und in der Praxis sogar erwünscht.

Diese neue angegebene Definition von $r_i \in \{0, 1\}$ bedeutet mit $u, v \in C^1$, dass als (verallgemeinerte) Ableitung von F^ε in (x, y) die Frechét-Ableitung in (x, y') genommen wird mit $y'_i = y_i + \delta_i$ und hinreichend kleinem $\delta_i > 0$. Dies entspricht in spezieller Weise dem Ansatz (3.13).

Bemerkung 6.3 Durch (6.24) darf dann im (ii) die Bedingung $b_i \neq 0$ weggelassen werden. Sie wird nämlich immer erfüllt, wenn wir (6.21) voraussetzen.

Bemerkung 6.4 (streng monotone $pNCP$ -Funktionen) Selbst die Beschränkung (6.21) kann vermieden werden, wenn wir NCP-Funktionen spezieller Klassen betrachten. Tatsächlich, die unreguläre Situation

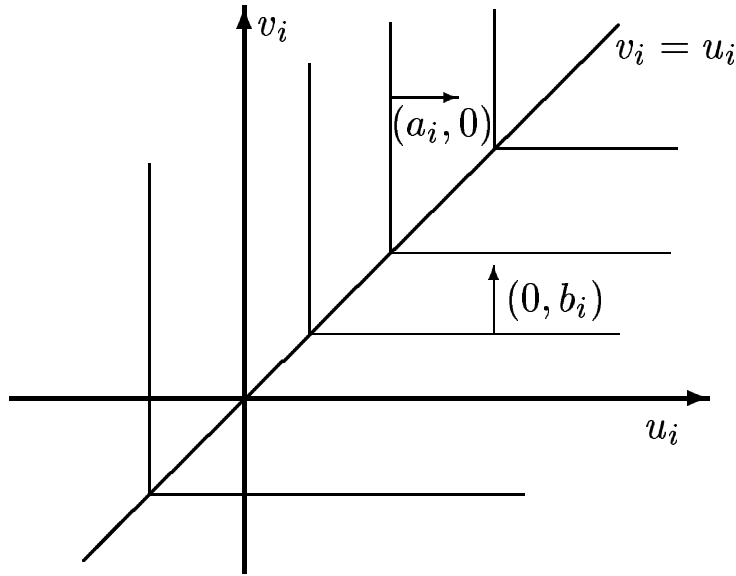
$$r_i = \frac{a_i}{a_i + b_i} \quad (6.25)$$

kann nur in den folgenden Fällen auftreten:

Fall 1. Sei $a_i b_i = 0$. Für streng monotone Funktionen (siehe Definition 6.2) heißt dies sofort, dass wir in der Lösung x^* des NCP sind. Es folgt nach der Definition 6.20 und der Diskussion danach (Seite 49), dass ε_i definiert bleibt: In der Situation I gilt $a_i(x^*) = 0$ (bzw. $b_i(x^*) = 0$), d.h. $u_i(x^*) > 0, v_i(x^*) = 0$ (bzw. $u_i(x^*) = 0, v_i(x^*) > 0$). Es folgt $y_i = u_i(x^*) - v_i(x^*) > 0$ (bzw. $y_i < 0$). Da $r \in R_T(y)$ impliziert das $r_i = 1$ (bzw. $r_i = 0$). Wir erhalten also immer $\varepsilon_i = 0$; In der Situation II ist ε_i beliebig.

Fall 2. Es kann passieren, dass zusammen mit (6.25) auch $0 < \frac{a_i}{a_i + b_i} < 1$ gilt; allerdings nur mit $y_i = 0$. In diesem Fall sieht die Gleichung (6.22) so aus:

$$\frac{b_i + \varepsilon_i a_i}{a_i + b_i} Lu_i(x, \xi) + \frac{\varepsilon_i b_i + a_i}{a_i + b_i} Lu_i(x, \xi) = 0$$

Abbildung 6.1: Niveaulinien von $G_{min} = \min\{u_i, v_i\}$

Es folgt also, dass es allgemein nicht möglich ist, daraus die Gleichung (6.12) zu bekommen. Deswegen müssen wir bei der Wahl von r_i aus der Menge $[0, 1]$ den Wert (6.25) (z.B. durch (6.24)) ausschließen. Aber in einem speziellen Fall, wenn $a_i = b_i$ gilt, kann man doch (6.12) erhalten, indem man einfach $\varepsilon_i = 2a_i - 1$ einsetzt. Dies läßt sich leicht sehen.

Weiter (siehe den nachfolgenden Abschnitt) wird genauer gezeigt wie der Satz 6.1 für die streng monotone NCP-Funktionen in Bezug auf die Restriktion (6.21) vereinfacht werden kann.

Bemerkung 6.5 (nicht streng monotone $pNCP$ -Funktionen) Zum gleichen Ergebniss (wie in Bemerkung 6.4) kommt man auch mit manchen anderen $pNCP$ -Funktionen. Wir betrachten z.B. die nicht streng monotone Funktion G_{min} .

Fall 1. Sei $a_i = 0$ und $y_i < 0$. Es folgt $r_i = 0$ und damit auch (6.25). Für die Funktion G_{min} bedeutet die Bedingung $a_i(x) = 0$ aber, dass der Punkt $(u_i(x), v_i(x))$ unter bzw. auf der Linie (Fläche) $v_i = u_i$ liegt (siehe Abbildung 6.1). Somit gilt $y_i = u_i(x) - v_i(x) \geq 0$.

Fall 2. Sei $b_i = 0$ und $y_i > 0$. Es folgt $r_i = 1$ und damit auch (6.25). Analog wie im Fall 1. erhalten wir, dass für die Funktion G_{min} diese Situation auch nicht zutrifft.

Fall 3. Sei $y_i = u_i(x) - v_i(x) = 0$. In dem Fall ist die Funktion G_{min} nicht differenzierbar. Bei der freien Wahl $(a_i, b_i) \in D^0 G_{min}(z_i(x))$ können wir allgemein z.B. immer $a_i = 0$, $b_i \neq 0$ festlegen. Sei also $a_i = 0$ und $y_i = 0$. Die Wahl $r_i = 0$ aus der Menge $[0, 1]$ führt dann zu (6.25). Wir vermeiden das, indem wir nach wie vor (6.24) in Anspruch nehmen.

Analog könnten auch andere nicht streng monotone $pNCP$ -Funktionen behandelt werden, falls sie die gleiche symmetrische (bezüglich der Flächen $v_i = u_i$, $\forall i$) Gestalt wie G_{min} haben.

6.3.2 Vereinfachung für streng monotone G aus $pNCP$

Wir betrachten nun streng monotone $G \in pNCP$ und lassen in System (6.19) nur spezielle $r \in R_T(y)$ zu.

Satz 6.2 Für festes $x \in \mathbb{R}^n$, seien a_i, b_i nach (6.11) erklärt und ξ erfülle die Gleichungen (6.12) mit streng monotonem $G \in pNCP$. Man setze $y = u(x) - v(x)$ und weiter $r_i = 1$ falls $y_i \geq 0$, $r_i = 0$ falls $y_i < 0$. Ferner sei $\varepsilon_i = 0$ wenn $a_i b_i = 0$, und andernfalls

$$\varepsilon_i = \frac{b_i}{a_i} \text{ wenn } r_i = 0, \quad \varepsilon_i = \frac{a_i}{b_i} \text{ wenn } r_i = 1. \quad (6.26)$$

Schließlich sei η definiert durch

$$\eta_i = \begin{cases} F_{n+i}^\varepsilon(x, y) - Dv_i(x)\xi, & \text{falls } r_i = 0 \\ F_i^\varepsilon(x, y) + Du_i(x)\xi, & \text{falls } r_i = 1. \end{cases}$$

Dann ist (ξ, η) eine Lösung des symmetrisch gestörten (verallgemeinerten) Kojima-Newton-Systems (6.19).

Beweis:

Wenn $a_i b_i \neq 0$, so sind die Argumente des Beweises (ii) zum letzten Satz voll anwendbar, weil ε_i und η_i wie dort definiert wurden und sowohl $(a_i + b_i)r_i \neq a_i$ als auch $r_i \neq \frac{1}{2}$ gilt.

Wenn $a_i b_i = 0$, so folgt aus der strengen Monotonie von G , dass $z_i(x) := (u_i(x), v_i(x))$ in $G^{-1}(0)$ liegt. Weiter gilt wegen $\varepsilon_i = 0$,

$$\begin{aligned} F_i^\varepsilon(x, y) &= F_i(x, y) = u_i(x) - y_i^+ = 0 \\ F_{n+i}^\varepsilon(x, y) &= F_{n+i}(x, y) = -v_i(x) - y_i^- = 0. \end{aligned}$$

Der Beweis folgt nun wegen $r_i \in R_T(y_i)$ auch mit den Argumenten zu Lemma 6.2 (ii), wir beweisen die Aussage zur besseren Übersicht direkt.

Sei $a_i = 0$. Dann kann nur $v_i = 0$ und $u_i \geq 0$ gelten. Damit ist $y_i = y_i^+ \geq 0$ und $r_i = 1$. Die zu erfüllenden Gleichungen aus (6.19) bedeuten so an der Stelle x ,

$$\begin{aligned} 0 &= u_i - y_i^+ + Du_i \xi - \eta_i = u_i - y_i^+ + Du_i \xi - (u_i(x) - y_i^+ + Du_i(x)\xi) = 0 \\ 0 &= -Dv_i \xi. \end{aligned}$$

Nach Voraussetzung (6.12) ist auch die zweite Gleichung wegen $v_i + Dv_i \xi = 0$ erfüllt. Analog folgt die Behauptung im Falle $b_i = 0$. ■

Die Bedeutung beider Sätze kann man folgendermaßen auffassen:

- Mit der Wahl einer (streng monotonen) $pNCP$ -Funktion ändern sich in der Gleichung (6.12) nur die Bewichtungen a_i bzw. b_i . In der Sprache von symmetrisch gestörten Kojima-Systemen (6.18) bedeutet das die entsprechende Wahl eines ε_i nach (6.20).
- Spezielle Lösungen des Systems (6.19) beschreiben nach Satz 6.2 durch die geeignete Wahl des Parameters $\varepsilon \geq 0$ die Lösungen der Systeme (6.12) mit allen möglichen streng monotonen $pNCP$ -Funktionen.

- Da wir den Parameter ε in (6.19) auch auf eine andere Weise variieren können (anderes als (6.20)) bzw. das System (6.15) auf eine andere Weise stören können (z.B. mit negativen ε_i) erhalten wir damit mehr Variationsmöglichkeiten als wenn wir einfach neue streng monotone $pNCP$ -Funktionen benutzen würden.
- Mit gewissen Einschränkungen (siehe Satz 6.1 und die Bemerkungen danach) sind diese Schlussfolgerungen auch für beliebige $pNCP$ -Funktionen G richtig.

Kapitel 7

Modelle für Variationsungleichungen

Mit Hilfe einer abgeschlossenen, konvexen Menge $C \subset \mathbb{R}^n$ und einer Funktion $p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ kann man eine Variationsungleichung (in der traditionellen Form) auf folgender Weise darstellen. Gesucht wird ein $x \in C$, so dass

$$p(x) \in N_C(x), \quad (7.1)$$

wobei

$$N_C(x) := \{\xi \in \mathbb{R}^n \mid \xi^T(c - x) \leq 0, \forall c \in C\}$$

Definition 7.1 Die Menge $N_C(x)$ heißt Normalenkegel von C im Punkt x . Wenn $x \notin C$, setzt man $N_C(x) := \emptyset$.

Habe die Menge C die folgende Gestalt

$$C = \{x \mid g_i(x) \leq 0, i = 1, \dots, m\}, \quad g_i \in C^1, \quad \forall i$$

Unter einer Constraint Qualification (nämlich genau unter der Abadie-Bedingung [19]) im Punkt x hat der Normalenkegel eine spezielle Form:

$$N_C(x) = \{y^+ Dg(x) \mid g(x) = y^-, y \in \mathbb{R}^m\},$$

d.h. (7.1) ist dann äquivalent zu

$$\begin{aligned} -p(x) + y^+ Dg(x) &= 0 \\ g(x) - y^- &= 0 \end{aligned} \quad (7.2)$$

Wie schon erwähnt wurde, ist die linke Seite in (7.2) ein weiteres Beispiel einer verallgemeinerten Kojima-Funktion (2.9). Das System kann man offenbar mittels einer $pNCP$ -Funktion $G(\cdot, \cdot)$ auch so darstellen:

$$\begin{aligned} -p(x) + y Dg(x) &= 0 \\ G(-g_i(x), y_i) &= 0, \quad \forall i = 1, \dots, m \end{aligned} \quad (7.3)$$

Die Gleichungen (7.2) und (7.3) haben zwar dieselben Lösungen, sind aber strukturell unterschiedlich. Dieser Unterschied verschwindet allerdings erneut, wenn wir das System (7.2) symmetrisch stören wie in (6.18) und dann auf die beiden Systeme das Newton-Verfahren anwenden.

Im weiteren betrachten wir diese Behauptung genauer und setzen generell $p \in C^1, g \in C^2$ voraus. Der allgemeinere Fall $p, Dg \in \text{loc}PC^1$ kann nach den Aussagen im Kapitel "Newton-Verfahren" (Seite 25) analog behandelt werden, nur müssen dann $Dp(x), D^2g(x)$ als Elemente aus $D^0p(x), D^0(Dg(x))$ interpretiert werden und die entsprechenden Voraussetzungen für alle solche Elemente erfüllt sein.

7.1 Einheitliche Darstellung des Newton–Verfahrens

Wir wollen zuerst sehen, wie die Newton–Gleichungen für (7.3) aussehen. Sei $(u, v) \in \mathbb{R}^{n+m}$ die gesuchte Newton–Richtung für einen Punkt (x, y) und $G := G(-g_i(x), y_i)$ zur Abkürzung, so ergibt sich

$$\begin{aligned} -p(x) + yDg(x) - Dp(x)u + yD^2g(x)u + vDg(x) &= 0 \\ G - G_sDg_i(x)u + G_tv_i &= 0 \quad \forall i \end{aligned}$$

Die partielle Ableitungen bezeichnen wir wie vorher $a_i := G_s(-g_i(x), y_i)$ und $b_i := G_t(-g_i(x), y_i)$. Dann folgt mit (6.4) weiter

$$\begin{aligned} -p(x) - Dp(x)u + yD^2g(x)u + \sum_i (y_i + v_i)Dg_i(x) &= 0 \\ -a_i[g_i(x) + Dg_i(x)u] + b_i[y_i + v_i] &= 0 \quad \forall i \end{aligned}$$

Wegen $DG \neq 0$ (siehe Lemma 6.1) erhalten wir aus der letzten Gleichung

$$\begin{aligned} y_i + v_i &= a_i/b_i[g_i(x) + Dg_i(x)u], \quad \text{falls } b_i \neq 0 \\ g_i(x) + Dg_i(x)u &= 0, \quad \text{falls } b_i = 0. \end{aligned} \quad (7.4)$$

Sei $K_1 := \{i \mid b_i = 0, i = \overline{1, m}\}$. Setzen wir $y_i + v_i$ in die erste Gleichung ein, so folgt

$$\begin{aligned} -p(x) - Dp(x)u + yD^2g(x)u + \\ \sum_{i \notin K_1} \frac{a_i}{b_i} [g_i(x) + Dg_i(x)u] Dg_i(x) + \sum_{i \in K_1} (y_i + v_i) Dg_i(x) &= 0 \end{aligned} \quad (7.5)$$

$$g_i(x) + Dg_i(x)u = 0 \quad \forall i \in K_1.$$

Analog betrachten wir nun das symmetrisch gestörte Kojima–System (7.2).

$$\begin{aligned} -p(x) + \sum_{i=1}^m (y_i^+ + \varepsilon_i y_i^-) Dg_i(x) &= 0 \\ g_i - (y_i^- + \varepsilon_i y_i^+) &= 0, \quad \forall i \end{aligned} \quad (7.6)$$

Differenziert man die linke Seite mittels Thibault–Ableitung TF (für CF ist es analog), bekommt man die folgenden Newton–Gleichungen

$$\begin{aligned} -p(x) + \sum_{i=1}^m (y_i^+ + \varepsilon_i y_i^-) Dg_i(x) - Dp(x)u + \\ + \sum_{i=1}^m (y_i^+ + \varepsilon_i y_i^-) D^2g_i(x)u + \sum_{i=1}^m (r_i + \varepsilon_i(1 - r_i)) v_i Dg_i(x) &= 0 \\ g_i(x) - (y_i^- + \varepsilon_i y_i^+) + Dg_i(x)u - (1 - r_i + \varepsilon_i r_i) v_i &= 0, \quad \forall i, \end{aligned}$$

wobei $r \in R_T(y)$ und (u, v) nach wie vor die gesuchten Newton–Richtungen für den Punkt (x, y) sind.

Führen wir die Menge $K_2 := \{i \mid 1 - r_i + \varepsilon_i r_i = 0, i = \overline{1, m}\}$ ein, folgt aus der zweiten Gleichung

$$v_i = \frac{g_i(x) + Dg_i(x)u - (y_i^- + \varepsilon_i y_i^+)}{1 - r_i + \varepsilon_i r_i}, \quad \text{falls } i \notin K_2$$

$$g_i(x) + Dg_i(x)u - (y_i^- + \varepsilon_i y_i^+) = 0, \quad \text{falls } i \in K_2$$

Setzen wir v_i in die erste Gleichung ein, erhalten wir

$$\begin{aligned}
& -p(x) - Dp(x)u + \sum_{i=1}^m (y_i^+ + \varepsilon_i y_i^-) D^2 g_i(x) u + \\
& + \sum_{i \notin K_2} (y_i^+ + \varepsilon_i y_i^- - \frac{r_i + \varepsilon_i(1-r_i)}{1-r_i + \varepsilon_i r_i} (y_i^- + \varepsilon_i y_i^+ - g_i(x) - Dg_i(x)u)) Dg_i(x) + \\
& + \sum_{i \in K_2} (y_i^+ + \varepsilon_i y_i^- + (r_i + \varepsilon_i(1-r_i))v_i) Dg_i(x) = 0 \\
& g_i(x) + Dg_i(x)u - (y_i^- + \varepsilon_i y_i^+) = 0, \quad \forall i \in K_2.
\end{aligned} \tag{7.7}$$

Da der Parameter ε_i frei variiert werden kann, setzen wir ihn diesmal so an:

$$\varepsilon_i = \frac{a_i(1-r_i) - b_i r_i}{b_i(1-r_i) - a_i r_i}, \quad \forall i \tag{7.8}$$

Im Vergleich mit (6.20) ist dies interessanterweise der reziproke Wert davon. Es gilt auch $\varepsilon_i \geq 0, \forall i$. Damit ε_i definiert bleibt, schließen wir natürlich den Fall

$$r_i = \frac{b_i}{a_i + b_i} \tag{7.9}$$

aus. Man sieht sofort die folgende Eigenschaften von ε_i :

$$\begin{aligned}
\varepsilon_i &= b_i/a_i, \quad \text{falls } y_i > 0 \text{ und } a_i \neq 0 \\
\varepsilon_i &= a_i/b_i, \quad \text{falls } y_i < 0 \text{ und } b_i \neq 0
\end{aligned} \tag{7.10}$$

Weiter überprüfen wir, genauso wie das schon für das NCP im vorigen Abschnitt gemacht wurde (siehe Seite 49), ob $\varepsilon \rightarrow 0$ gilt, wenn eine Folge $(x, y)_\varepsilon$ der Lösungen der Systeme (7.6) gegen eine Lösung der Variationsungleichung (7.1) konvergiert.

Tatsächlich, sei ε wie in (7.8) angegeben. Wir betrachten eine Folge $(x, y) \rightarrow (x^*, y^*)$ nahe an der Lösung (x^*, y^*) der Aufgabe (7.2) und untersuchen verschiedene Fälle:

Situation I (strikte Komplementarität):

Sei $\sigma_i^* := (-g_i(x^*), y_i^*) \neq (0, 0)$.

Fall 1. $-g_i(x^*) > 0$

Nach dem Lemma 6.1 und (6.3) folgt $a_i(\sigma_i^*) = 0$ und $b_i(\sigma_i^*) > 0$. D.h. für (x, y) nahe (x^*, y^*) gilt: $a_i(x, y) \rightarrow 0, b_i(x, y) > 0$. Außerdem gilt $g_i(x) < 0$. Dies liefert uns $y_i < 0$, weil $g_i(x) = y_i^- + \varepsilon_i y_i^+$. Nach (7.10) erhalten wir letztendlich $\varepsilon_i = a_i(x, y)/b_i(x, y) \rightarrow 0$.

Fall 2. $y_i^* > 0$

Das Lemma 6.1 und (6.3) liefern uns $b_i(\sigma_i^*) = 0$ und $a_i(\sigma_i^*) > 0$. Nach (7.10) folgt also auch $\varepsilon_i = b_i(x, y)/a_i(x, y) \rightarrow 0$.

Situation II : Sei $\sigma_i^* = (0, 0)$. Daraus folgt $y_i \rightarrow 0$ und $g_i(x) \rightarrow 0$. Das garantiert uns, dass wir zur Lösung des System (7.2) unabhängig vom Verhalten des Parameters ε_i (vorausgesetzt er bleibt beschränkt) kommen.

Man sieht also, dass die Wahl von ε nach (7.8) legitim ist. Weiter betrachten wir den regulären Fall

$$r_i \neq \frac{b_i}{a_i + b_i} \tag{7.11}$$

von (7.8) genauer.

Man stellt hier fest, dass $K_1 \subset K_2$ gilt. Tatsächlich, wenn $b_i = 0$ ist, erhält man $\varepsilon_i = (r_i - 1)/r_i$, und damit auch $1 - r_i + \varepsilon_i r_i = 0$, also $i \in K_2$.

Wir diskutieren als nächstes die Möglichkeiten von b_i, y_i und treffen Vereinbarungen für die Wahl von r_i , wenn $y_i = 0$.

Sei $i \in K_1$, d.h. $b_i = 0$. Dann kann der Fall $y_i < 0$ nie auftreten, denn sonst wären wir wegen $r_i = 0$ im nicht regulären Fall (7.9). Wenn $y_i > 0$ gilt, folgt $r_i = 1$ und $\varepsilon_i = 0$. Wenn $y_i = 0$ ist, so setzen wir (da wir dann freie Wahl $r_i \in [0, 1]$ haben) wie gewohnt (6.24) $r_i = 1$.

Sei nun $i \in K_2 \setminus K_1$. Wenn wir ε_i hier auch wie in (7.8) bestimmen, folgt sofort $r_i = 1/2$. Dies impliziert wegen $r \in R_T(y)$ aber $y_i = 0$. Unsere Vereinbarung (6.24) schließt den Wert $r_i = 1/2$ einfach aus.

Mit (6.24) gilt im regulären Fall (7.11) somit in der Tat

$$K_1 = K_2.$$

Sei (7.8) für alle $i = 1, \dots, m$ definiert. Setzen wir jetzt den Parameter in das System (7.7) ein. Mit Berücksichtigung der Überlegungen oben erhalten wir nun das folgende Newton-System

$$\begin{aligned} & -p(x) - Dp(x)u + \sum_{i=1}^m (y_i^+ + \varepsilon_i y_i^-) D^2 g_i(x)u + \\ & + \sum_{i \notin K_1} \frac{a_i}{b_i} [g_i(x) + Dg_i(x)u] Dg_i(x) + \sum_{i \in K_1} (y_i + v_i) Dg_i(x) = 0 \end{aligned} \quad (7.12)$$

$$g_i(x) + Dg_i(x)u = 0, \quad \forall i \in K_1$$

Wenn man nun das System mit den Gleichungen (7.5) vergleicht, sieht man, dass es einen einzigen Unterschied im dritten Term der ersten Gleichung gibt: Anstatt einfach y_i steht hier $y_i^+ + \varepsilon_i y_i^-$. Dieser Unterschied verschwindet allerdings, wenn z.B. $y_i \geq 0$ ist.

Somit sind wir zu der folgenden Aussage gekommen.

Satz 7.1 *Das Paar (u, v) ist eine Newton-Richtung in Punkt (x, \hat{y}) der Aufgabe (7.3) mit einer pNCP-Funktion G genau dann, wenn (u, v) die Newton-Richtung für das symmetrisch gestörte Kojima-System (7.6) in Punkt (x, y) ist, sobald ε_i nach (7.8) definiert bleibt und $\varepsilon_i \neq 0$ für $\hat{y}_i < 0$. Dabei gilt:*

$$y_i = \begin{cases} \hat{y}_i, & \hat{y}_i \geq 0 \\ \frac{1}{\varepsilon_i} \hat{y}_i, & \hat{y}_i < 0 \end{cases}$$

oder

$$\hat{y}_i = y_i^+ + \varepsilon_i y_i^-$$

für alle $i = 1, \dots, m$

Beweis:

Wir vergleichen (7.12) und (7.5). In der letzten Summe der ersten Gleichung von (7.12) gilt $b_i = 0$ ($i \in K_1$) und wenn $y_i < 0$ ist, sind wir dann im nicht regulären Fall (7.9). Der Parameter ε_i bleibt also undefiniert. Nach der Voraussetzung des Satzes muß hier also $y_i \geq 0$ gelten. Hiermit ist die Beziehung $y_i = y_i^+ + \varepsilon_i y_i^-$, $\forall i \in K_1$ trivialerweise erfüllt und der Satz ist bewiesen. ■

Wie in der Bemerkung 6.4 betrachten wir die nicht regulären Situationen (7.9) genauer:

Bemerkung 7.1 (streng monotone $pNCP$ –Funktionen)

Fall 1. Sei $a_i b_i = 0$. Nach Definition 6.2 folgt, dass wir in einer Nullstelle $(g_i(x^*), y_i^*)$ der NCP-Funktion G sind. Nach (7.8) und Diskussion danach (Seite 57) bleibt das Parameter ε_i dann definiert: In der Situation I haben wir $a_i = 0$ (bzw. $b_i = 0$) $\implies y_i^* < 0$ (bzw. $y_i^* > 0$) $\implies r_i = 0$ (bzw. $r_i = 1$). Damit ist in dem Fall immer $\varepsilon_i = 0$ bleibt; In der Situation II darf ε_i beliebig gewählt werden.

Fall 2. Gelte nun $0 < \frac{b_i}{a_i + b_i} < 1$ und (7.9). Es kann nur dann passieren, wenn $y_i = 0$ ist. Dieser Fall wird leicht ausgeschlossen, wenn wir nach wie vor (6.24) ansetzen. Man braucht allerdings auch das nicht zu machen, falls $a_i = b_i$ gilt (ε_i darf dann beliebig sein). Dies lässt sich leicht nachprüfen.

Bemerkung 7.2 (nicht streng monotone $pNCP$ –Funktionen) Wir betrachten wieder die einfachste nicht streng monotone $pNCP$ –Funktion $G = G_{min}$. Wie vorher stellen wir fest, dass für sie die Bedingung (7.11) auch immer erfüllt bleibt. Tatsächlich:

Fall 1. Sei $a_i = 0$ bzw. $b_i = 0$. Dies liefert uns zusammen mit $y_i > 0$ bzw. $y_i < 0$ den nicht regulären Fall (7.9), da dann $r_i = 1$ bzw. $r_i = 0$ wegen $r \in R_T(y)$ gilt. Nach der Abbildung 6.1 erkennt man, dass wir uns unter bzw. über der Linien $v_i = u_i$ befinden. D.h. für $-g_i(x)$ (steht nun für u_i) und y_i (steht nun für v_i) einerseits

$$-g_i(x) > y_i > 0 \quad \text{bzw.} \quad -g_i(x) < y_i < 0,$$

andererseits

$$g_i(x) = y_i^- + \varepsilon_i y_i^+ \tag{7.13}$$

gilt. Dies widerspricht aber unseren Annahmen $y_i > 0$ bzw. $y_i < 0$, denn nach (7.13) folgt $y_i < 0$ bzw. $y_i > 0$. Damit ist der Fall ausgeschlossen.

Fall 2. Die Situation, wenn wir auf der Linie $v_i = u_i$ – sprich $-g_i(x) = y_i$ sind, kommt wegen (7.13) nur dann vor, falls $-g_i(x) = y_i = 0$ gilt. In dem Punkt ist die Funktion G_{min} aber nicht (Frechet–) differenzierbar. Deshalb können wir bei der freien Wahl $(a_i, b_i) \in D^0 G_{min}(0, 0)$ allgemein z.B. immer $a_i \neq 0$, $b_i = 0$ festlegen. Sei also $b_i = 0$ und $y_i = 0$. Die Wahl $r_i = 0$ aus der Menge $[0, 1]$ führt dann zum (7.9). Dies wird aber auch nicht auftreten, wenn wir nach wie vor (6.24) ansetzen.

Für die andere symmetrische (bzg. der Fläche $v_i = u_i$) $pNCP$ –Funktionen kann man analog wie in der vorigen Bemerkung zum gleichen Ergebnis kommen.

Analog zum Satz 6.1 haben wir in diesem Abschnitt also gesehen, dass beim Newton–Verfahren auch für die Variationsungleichungen der NCP–Ansatz (7.3) sehr leicht von der symmetrisch gestörten Kojima–System (7.6) modelliert werden kann. Mit anderen Worten: Egal mit welcher Beschreibung der Variationsaufgabe wir uns beschäftigen, mit (7.3) oder mit (7.6) — in der Tat wird durch die angegebene Wahl von ε_i (7.8) das gleiche System gelöst.

Mit den gestörten Kojima–Systeme haben wir also einen größeren Raum zum Variieren im Vergleich zu NCP–Ansätzen aller Art. Solche Kojima–Systeme sind damit ein nicht weniger sinnvolles Instrument zur Behandlung auch allgemeinerer Probleme, die ein NCP als Teilproblem enthalten. Ihr Vorteil besteht in den leicht interpretierbaren Nullstellen und den direkten Möglichkeiten der Lösungsabschätzung.

Im nächsten Abschnitt wird darüberhinaus gezeigt, dass sich auch die Regularitätsbedingungen fürs Newton–Verfahren zur Aufgabe (7.3) mit allen möglichen $pNCP$ –Funktion–Ansätzen über die entsprechende Regularitätsbedingung des Kojima–System (7.2) einheitlich beschreiben lassen.

7.2 Verlauf der Verfahren

7.2.1 Gradienten–Verfahren und Monotonie

Um das System (7.3) zu lösen, macht die Benutzung der Newton–Verfahren nur dann einen Sinn, wenn wir schon hinreichend nahe an der Lösung sind. Um das zu gewährleisten kann zum Beispiel zuerst ein Abstiegsverfahren erster Ordnung angewandt werden. So, wie schon im Abschnitt 6.1 erwähnt wurde, versucht man dann anstatt des Gleichungssystems (7.3) ein Minimierungsproblem zu lösen.

$$\min_{x,y} Q(x,y), \quad (7.14)$$

wobei

$$Q(x,y) := \frac{1}{2} \sum_{k=1}^n h_k^2(x,y) + \frac{1}{2} \sum_{i=1}^m h_i^2(x,y) \quad (7.15)$$

eine Hilfszielfunktion (Merit-Funktion) ist und

$$\begin{aligned} h_k(x,y) &:= -p_k(x) + [yDg(x)]_k, \quad k = 1, \dots, n \\ h_i(x,y) &:= G(-g_i(x), y_i), \quad i = 1, \dots, m \end{aligned}$$

Der Term $[yDg(x)]_k$ ist hier das k -te Element des Vektors. Damit das Verfahren funktioniert, werden weiter die folgende Eigenschaften gefordert.

Erstens möge die Funktion $Q(x,y)$ aus der Klasse C^1 sein. Das wird automatisch erreicht, wenn wir annehmen, dass

$$g \in C^2, \quad p \in C^1, \quad (7.16)$$

und G so wählen, dass

$$\Theta^1(G) \cup G^{-1}(0) = \mathbb{R}^2 \quad (7.17)$$

Zweitens sollte aus $DQ(x,y) = 0$ folgen, dass (x,y) eine Lösung von (7.3) ist, d.h. $Q(x,y) = 0$.

Im allgemeinen gilt dies nicht, aber unter strenger Monotonie von G und einer Definitheitsforderung ist diese Implikation gesichert.

Lemma 7.1 *Sei (7.16) und (7.17) erfüllt und sei G streng monoton. Außerdem sei die Matrix $-Dp(x) + yD^2g(x)$ positiv definit. Dann ist $DQ(x,y) = 0$ hinreichend für $Q(x,y) = 0$.*

Beweis:

Der Einfachheit halber sei $F_1 := \{h_k(x,y)\}_{k=1}^n$, $F_2 := \{h_i(x,y)\}_{i=1}^m$ und $\sigma_i := (-g_i(x), y_i)$. Das heißt also $F_1 \in \mathbb{R}^n$, $F_2 \in \mathbb{R}^m$ und $D_x F_1 = -Dp(x) + yD^2g(x)$. Aus $DQ(x,y) = 0$ folgt, dass die partiellen Ableitungen von $Q(x,y)$ verschwinden

$$\begin{aligned} D_x Q(x,y) &= F_1(D_x F_1) - \sum_{i=1}^m F_{2i} G_s(\sigma_i) Dg_i(x) = 0 \\ D_{y_i} Q(x,y) &= Dg_i(x) F_1 + F_{2i} G_t(\sigma_j) = 0 \quad i = 1, \dots, m \end{aligned}$$

Multipliziert man die erste Gleichung mit F_1 , erhält man

$$\begin{aligned} F_1(D_x F_1) F_1 - \sum_{i=1}^m F_{2i} G_s(\sigma_i) Dg_i(x) F_1 &= 0 \\ Dg_i(x) F_1 = -F_{2i} G_t(\sigma_i) \quad i = 1, \dots, m \end{aligned}$$

Nun setzen wir die zweite Reihe von Gleichungen in die erste ein:

$$F_1(D_x F_1)F_1 + \sum_{i=1}^m F_{2i}^2 G_s(\sigma_i)G_t(\sigma_i) = 0.$$

Unter der Voraussetzungen des Lemmas kann diese nur dann erfüllt werden, wenn $F_1(x, y) = 0$ und $F_2(x, y) = 0$. Also $Q(x, y) = 0$. ■

Bemerkung 7.3 Auch wenn die NCP-Funktion G die Glattheitsbedingung (7.17) nicht erfüllt, kommt man zur gleichen Schlußfolgerung. Man muß einfach die partiellen Ableitungen $(G_s(\sigma_i), G_t(\sigma_i))$ durch die Paare $(a_i, b_i) \in D^0 G(\sigma_i)$ ersetzen.

7.2.2 Newton-Verfahren

Wenn wir nach dem Abstiegsverfahren hinreichend nahe an der Lösung sind, können wir das Newton-Verfahren für (7.3) ansetzen. Sei $p \in C^1$, $g \in C^2$ und erfülle G (6.9). Das bedeutet, dass es Punkte geben kann, wo die partiellen Ableitungen G_s und G_t nicht existieren. Im Rahmen der D^0 -Ableitung können die Ableitungen dann durch beliebige Limites aus der Menge

$$D^0 G(\sigma) = \left\{ \lim \begin{pmatrix} G_s(\sigma^k) \\ G_t(\sigma^k) \end{pmatrix} \mid \sigma^k \rightarrow \sigma, \sigma^k \in \Theta^1(G) \right\} \quad (7.18)$$

ersetzt werden. Zur Abkürzung werden wir allerdings auch weiter G_s und G_t schreiben.

Die Newton-Gleichung für die Richtung (u, v) sieht dann so aus:

$$\begin{aligned} -Dp(x)u + yD^2g(x)u + vDg(x) &= -F_1 \\ -G_s(\sigma_i)Dg_j(x)u + G_t(\sigma_i)v_i &= -G(\sigma_i), \quad j = 1, \dots, m \end{aligned}$$

wobei die Bezeichnungen σ_i und F_1 dieselbe Bedeutung wie in Lemma 7.1 haben.

Lemma 7.2 Sei $G_t(\sigma_i) > 0$, $\forall i$. Ist die Matrix $M(x, y) := -Dp(x) + yD^2g(x)$ positiv definit, so auch die Newton-Matrix (bzw. Matrizen[†]) $P(x, y)$:

$$\begin{pmatrix} M(x, y) & Dg_1(x) & \cdots & Dg_i(x) & \cdots & Dg_m(x) \\ -G_s(\sigma_1)Dg_1(x)^T & G_t(\sigma_1) & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -G_s(\sigma_i)Dg_i(x)^T & 0 & \cdots & G_t(\sigma_i) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -G_s(\sigma_m)Dg_m(x)^T & 0 & \cdots & 0 & \cdots & G_t(\sigma_m) \end{pmatrix} \quad (7.19)$$

Beweis:

Wir wollen also zeigen, dass

$$(\alpha \ \beta)P(x, y) \begin{pmatrix} \alpha \\ \beta \end{pmatrix} > 0, \quad \forall \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \neq 0$$

[†] falls die Menge (7.18) mehrelementig ist

und $\alpha \in \mathbb{R}^n$, $\beta \in \mathbb{R}^m$. Explizit sieht diese quadratische Form so aus

$$\alpha^T M(x, y) \alpha + \sum_{i=1}^m [\beta_i (1 - G_s(\sigma_i)) Dg_i(x)^T \alpha + \beta_i^2 G_t(\sigma_i)] \quad (7.20)$$

Nach der Voraussetzung des Lemmas reicht es zu zeigen, dass jeder Term in der letzten Summe für alle (α, β) nicht negativ ist.

Nun fixieren wir ein beliebiges α und betrachten jeden solchen Term als eine Funktion von β_i . Da $G_t(\sigma_i) > 0$, nimmt sie das Minimum an. Und zwar genau dann, wenn gilt

$$(1 - G_s(\sigma_i)) Dg_i(x)^T \alpha + 2\beta_i G_t(\sigma_i) = 0$$

D.h. der extremale Punkt hat die Form

$$\beta_i^* = -\frac{1}{2} \frac{(1 - G_s(\sigma_i))}{G_t(\sigma_i)} Dg_i(x)^T \alpha$$

Wir können jetzt jeden Term mittels des Minimalwertes im Punkt β_i^* nach unten abschätzen. Setzen wir also β_i^* in den entsprechenden i -ten Term der Summe ein, erhalten wir

$$\begin{aligned} & -\frac{1}{2} \frac{(1 - G_s(\sigma_i))}{G_t(\sigma_i)} (1 - G_s(\sigma_i)) (Dg_i(x)^T \alpha)^2 + \\ & + \frac{(1 - G_s(\sigma_i))^2}{G_t(\sigma_i)} (Dg_i(x)^T \alpha)^2 = \frac{1}{2} \frac{(1 - G_s(\sigma_i))^2}{G_t(\sigma_i)} (Dg_i(x)^T \alpha)^2 \end{aligned} \quad (7.21)$$

Man sieht, dass diese Funktion für alle α nicht negativ bleibt. ■

Es ist möglich, die Voraussetzungen des Lemmas ein wenig abzuschwächen.

Satz 7.2 Sei $G_t(\sigma_i) > 0$ und $G_s(\sigma_i) \neq 1$, $\forall i$. Sei die Matrix $M(x, y)$ positiv semidefinit und auf dem Unterraum $\{\alpha \in \mathbb{R}^n \mid Dg(x)\alpha = 0\}$ positiv definit. Dann folgt, dass die Newton-Matrizen $P(x, y)$ auch positiv definit sind.

Beweis:

Die quadratische Form (7.20) ist genau dann positiv definit, wenn stets einer der beiden (nichtnegativen) Terme in (7.20) für $(\alpha, \beta) \neq 0$ positiv bleibt. Der Minimalwert (7.21) für den i -ten Term der Summe sagt uns, dass das unter der Voraussetzungen des Satzes gesichert wird. ■

Der folgende Satz zeigt, dass das Newton-Verfahren auch ohne die positive Definitheit (eine zu starke Voraussetzung) für die Matrix M funktionieren kann.

Zunächst definieren wir eine Matrix Γ :

$$\Gamma := \begin{pmatrix} \gamma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \gamma_m \end{pmatrix},$$

wobei $\gamma_i = G_s(\sigma_i)/G_t(\sigma_i)$, $\forall i$.

Satz 7.3 Die Newton–Matrizen $P(x, y)$ sind regulär genau dann, wenn die Matrix $M(x, y) + Dg(x)^T \Gamma Dg(x)$ regulär ist, sofern $G_t(\sigma_i) \neq 0$, $\forall i$ gilt.

Beweis:

Sei $P(x, y)$ singular. Betrachten wir das Gleichungssystem

$$P(x, y) \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0,$$

wobei $(\alpha, \beta) \neq 0$ und $\alpha \in \mathbb{R}^n$, $\beta \in \mathbb{R}^m$.

Explizit sieht es so aus:

$$\begin{aligned} M(x, y)\alpha + Dg(x)^T \beta &= 0 \\ -G_s(\sigma_i)Dg_i(x)^T \alpha + G_t(\sigma_i)\beta_i &= 0, \quad \forall i = 1, \dots, m \end{aligned} \quad (7.22)$$

Da $G_t(\sigma_i) \neq 0$ ist, lassen sich die letzten m Gleichungen dann so umschreiben

$$\beta_i = [G_s(\sigma_i)/G_t(\sigma_i)]Dg_i(x)^T \alpha, \quad \forall i = 1, \dots, m$$

Es folgt

$$\beta = \Gamma Dg(x)\alpha.$$

Setzen wir β in der ersten Gleichung (7.22) ein, erhalten wir

$$(M(x, y) + Dg(x)^T \Gamma Dg(x))\alpha = 0.$$

Da $\alpha \neq 0$ ist, gilt dies nur dann, wenn die Matrix $M(x, y) + Dg(x)^T \Gamma Dg(x)$ singular ist. Wenn man nun die Schritte in umgekehrter Richtung (von unten nach oben) verfolgt, kommt man zu der Aussage des Satzes. ■

Bemerkung 7.4 Es ist leicht zu sehen, dass die Voraussetzungen des Lemmas 7.2 stärker als in Satz 7.3 sind. Die Matrix $M(x, y) + Dg(x)^T \Gamma Dg(x)$ ist nämlich positiv definit (bzw. semidefinit), wenn $M(x, y)$ positiv definit (bzw. semidefinit) ist. Dies folgt einfach aus der Semidefinitheit von Matrix $Dg(x)^T \Gamma Dg(x)$ (weil $\gamma_i \geq 0$, $\forall i$).

Regularität der Newton–Matrizen des Kojima–Systems

Wir kehren jetzt zu dem Kojima–System (7.2) zurück. Die Thibault–Ableitung der entsprechenden Kojima–Funktion ist mit der Menge der folgenden Matrizen $P^+(x, y)$ identifizierbar

$$\begin{pmatrix} M^+(x, y) & r_1 Dg_1(x) & \cdots & r_i Dg_i(x) & \cdots & r_m Dg_m(x) \\ Dg_1(x)^T & -(1 - r_1) & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ Dg_i(x)^T & 0 & \cdots & -(1 - r_i) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ Dg_m(x)^T & 0 & \cdots & 0 & \cdots & -(1 - r_m) \end{pmatrix} \quad (7.23)$$

wobei $r \in R_T(y)$ und $M^+(x, y) = -Dp(x) + y^+ D^2 g(x)$.

Lemma 7.3 Für beliebiges $r \in \mathbb{R}^m$ ist die Matrix (7.23) genau dann regulär, wenn die Matrix $Q^+(x, y)$:

$$\begin{pmatrix} M^+(x, y) & Dg_1(x) & \cdots & Dg_i(x) & \cdots & Dg_m(x) \\ r_1 Dg_1(x)^T & -(1-r_1) & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ r_i Dg_i(x)^T & 0 & \cdots & -(1-r_i) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ r_m Dg_m(x)^T & 0 & \cdots & 0 & \cdots & -(1-r_m) \end{pmatrix} \quad (7.24)$$

regulär ist.

Beweis:

Wir betrachten zwei Fälle. **Fall 1:** $r_i \neq 0$ und $r_i \neq 1$, $\forall i = 1, \dots, m$.

Die Gleichung $P^+(\alpha, \beta)^T = 0$ mit $(\alpha, \beta) \in \mathbb{R}^{n+m}$ sieht dann folgendermaßen aus

$$\begin{cases} M^+(x, y)\alpha + \sum_{i=1}^m r_i Dg_i(x)\beta_i = 0 \\ Dg_i(x)^T \alpha - (1-r_i)\beta_i = 0, \quad \forall i = 1, \dots, m \end{cases}$$

Weiter erhält man

$$\begin{cases} M^+(x, y)\alpha + \sum_{i=1}^m \frac{r_i}{1-r_i} Dg_i(x) Dg_i(x)^T \alpha = 0 \\ \beta_i = \frac{1}{1-r_i} Dg_i(x)^T \alpha, \quad \forall i = 1, \dots, m \end{cases} \quad (7.25)$$

Andererseits bekommen wir aus der Gleichung $Q^+(\alpha, \beta)^T = 0$:

$$\begin{cases} M^+(x, y)\alpha + \sum_{i=1}^m Dg_i(x)\beta_i = 0 \\ r_i Dg_i(x)^T \alpha - (1-r_i)\beta_i = 0, \quad \forall i = 1, \dots, m \end{cases}$$

oder

$$\begin{cases} M^+(x, y)\alpha + \sum_{i=1}^m \frac{r_i}{1-r_i} Dg_i(x) Dg_i(x)^T \alpha = 0 \\ \beta_i = \frac{r_i}{1-r_i} Dg_i(x)^T \alpha, \quad \forall i = 1, \dots, m \end{cases} \quad (7.26)$$

Da die Koeffizienten in der ersten Gleichung jeweils identisch sind, ist nun leicht zu sehen: wenn das System (7.25) trivial lösbar ist, so auch das System (7.26), und umgekehrt.

Fall 2: $\exists j$ bzw. $\exists k$ so, dass $r_j = 0$ bzw. $r_k = 1$.

So wie im Fall 1 betrachten wir die beiden Gleichungen: $P^+(\alpha, \beta)^T = 0$ und $Q^+(\alpha, \beta)^T = 0$. Nach der trivialen Transformation sieht das erste System so aus:

$$\begin{cases} M^+(x, y)\alpha + \sum_{i=1}^m \sum_{i \neq j, i \neq k} \frac{r_i}{1-r_i} Dg_i(x) Dg_i(x)^T \alpha + \sum_k Dg_k(x)\beta_k = 0 \\ \beta_i = \frac{1}{1-r_i} Dg_i(x)^T \alpha, \quad i \neq j, \quad i \neq k \\ \beta_j = D^T g_j(x)\alpha \\ Dg_k(x)^T \alpha = 0 \end{cases}$$

und das zweite so:

$$\begin{cases} M^+(x, y)\alpha + \sum_{i=1}^m \sum_{i \neq j, i \neq k} \frac{r_i}{1-r_i} Dg_i(x) Dg_i(x)^T \alpha + \sum_k Dg_k(x)\beta_k = 0 \\ \beta_i = \frac{r_i}{1-r_i} Dg_i(x)^T \alpha, \quad i \neq j, \quad i \neq k \\ \beta_j = 0 \\ Dg_k(x)^T \alpha = 0 \end{cases}$$

Hier ist, wie auch im Fall 1, leicht zu sehen, dass die ersten (entscheidenden) Gleichungen der beiden Systeme identisch sind. So folgt erneut: Wenn ein System trivial lösbar ist, so ist es das andere auch. ■

Folgerung 7.1 Die Kojima-Funktion aus (7.2) ist an der Stelle (x, y) streng regulär genau dann, wenn die entsprechenden Matrizen $Q^+(x, y)$ (siehe (7.24)) für alle $r \in R_T(y)$ regulär sind.

Beweis:

Man wendet den Satz 3.1 an. ■

Der folgende Satz ermittelt unter anderem, wie die Regularitätsbedingung für das Newton-Verfahren der Aufgabe (7.3) mit der strengen Regularität der Kojima-Funktion aus (7.2) zusammenhängt.

Satz 7.4 Es sei (x^*, y^*) eine Lösung von (7.3).

- (i) Wenn die Matrizen $P(x^*, y^*)$ (siehe (7.19)) für eine pNCP-Funktion $G(s, t)$ regulär sind, so bleiben sie regulär auch für die pNCP-Funktion: $G_{\min}(s, t) = \min\{s, t\}$.
- (ii) Die Matrizen $P(x^*, y^*)$ sind regulär, wenn die Kojima-Funktion aus (7.2) an der Stelle (x^*, y^*) streng regulär ist.
- (iii) Die Regularität der Matrizen $P(x^*, y^*)$ ist äquivalent mit der strengen Regularität an der Stelle (x^*, y^*) der Kojima-Funktion aus (7.2), wenn die Einheitsvektoren in \mathbb{R}^2 mittels eines Bogens in $\Theta^1(G)$ verbunden werden können.

Beweis:

Nach dem Lemma 6.1 folgt $\max\{a_i, b_i\} \geq p$ für ein $p > 0$. Man stellt leicht fest, dass die Matrizen $P(x, y)$ (siehe (7.19)) genau dann regulär sind, wenn dasselbe für die folgenden Matrizen gilt:

$$Q(x, y) = \begin{pmatrix} M(x, y) & Dg_1(x) & \cdots & Dg_i(x) & \cdots & Dg_m(x) \\ \lambda_1 Dg_1(x)^T & -(1 - \lambda_1) & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \lambda_i Dg_i(x)^T & 0 & \cdots & -(1 - \lambda_i) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \lambda_m Dg_m(x)^T & 0 & \cdots & 0 & \cdots & -(1 - \lambda_m) \end{pmatrix}$$

wobei $\lambda_i = a_i/(a_i + b_i)$, $1 - \lambda_i = b_i/(a_i + b_i)$ und $(a_i, b_i) \in D^0G(-g_i(x), y_i)$. Das bedeutet, dass die Koeffizienten λ_i eine Menge $S_i(x, y) \subset [0, 1]$ bilden. Je kleiner diese Menge ist, desto schwächer ist die Bedingung. Setzen wir nun $\sigma_i^* = (-g_i(x^*), y^*)$ an, und betrachten die möglichen Situationen. Wenn $-g_i(x^*) > 0$ gilt, dann folgt $y^* = 0$ und nach dem Lemma 6.1 mit (6.3) erhalten wir $a_i = 0, b_i > 0$ im Punkt σ_i^* ; d.h. $\lambda_i = 0$. Analog, wenn $y^* > 0$ gilt, so folgt $\lambda_i = 1$.

Sei $\sigma_i^* = (0, 0)$. Jetzt nimmt das Paar (a_i, b_i) beliebige Werte aus der Menge $\text{cl } DG(\Theta^1(G))$ an (siehe (6.5)), und

$$S_i(x^*, y^*) = \left\{ \frac{a_i}{a_i + b_i} \mid (a_i, b_i) \in \text{cl } DG(\Theta^1(G)) \right\} \quad (7.27)$$

Im "kleinsten" Fall enthält $S_i(x^*, y^*)$ nur 0 und 1. Das ist gerade der Fall, wenn die NCP-Funktion $G = G_{\min}$ ist. Im "größten" Fall gehört das ganze Intervall $[0, 1]$ zu $S_i(x^*, y^*)$. Hier

stimmt die Regularität von allen Matrizen $Q(x^*, y^*)$ nach der Folgerung 7.1 mit der strengen Regularität an der Stelle (x^*, y^*) der Kojima-Funktion aus (7.2) überein, da $(y^*)^+ = y^*$ (bzw. $M^+(x^*, y^*) = M(x^*, y^*)$) – sprich $Q^+(x^*, y^*) = Q(x^*, y^*)$ gilt. Damit sind die Behauptungen (i) und (ii) des Satzes bewiesen. Wenn wir nun einen (stetigen) Bogen in $\Theta^1(G)$ haben, der zwei Einheitsvektoren e^1 und e^2 verbindet, so ist die Menge $S_i(x^*, y^*)$ zusammenhängend und enthält 0 und 1 (siehe(7.27)). Das bedeutet also, dass wir im "größten" Fall sind. ■

Bemerkung 7.5 Die Behauptung (ii) gilt auch für beliebige Punkte (x, y) , sofern $y \geq 0$ ist.

Mit dem Satz haben wir also das Folgende hergeleitet.

- Die NCP-Funktion G_{min} taugt zwar für das Gradienten-Verfahren nicht (weil sie keine streng monotone NCP-Funktion ist), sie ist aber für das Newton-Verfahren die "einfachste".
- Die Regularität für das Newton-Verfahren von Aufgabe (7.3) kann allgemein und unabhängig von dem konkreten Aussehen der NCP-Funktion festgestellt werden.

Literaturverzeichnis

- [1] J.-P. Aubin and I. Ekeland. *Applied Nonlinear Analysis*. Wiley, New York, 1984.
- [2] F.H. Clarke. On the inverse function theorem. *Pacific Journal of Mathematics*, 64:97–102, 1976.
- [3] F.H. Clarke. *Optimization and Nonsmooth Analysis*. Wiley, New York, 1983.
- [4] A. Dontchev and R.T. Rockafellar. Characterizations of strong regularity for variational inequalities over polyhedral convex sets. *SIAM Journal on Optimization*, 6:1087–1105, 1996.
- [5] A. Fischer. Solutions of monotone complementarity problems with locally Lipschitzian functions. *Mathematical Programming, Series B*, 76:513–532, 1997.
- [6] P. Fusek. Isolated zeros of Lipschitzian metrically regular \mathbb{R}^n functions. *Optimization*, 49:425–446, 2001.
- [7] P. Fusek. *Eigenschaften pseudo-regulärer Funktionen und einige Anwendungen auf Optimierungsaufgaben*. Dissertation, Fachbereich Mathematik. Humboldt–Universität zu Berlin, Berlin, Februar 1999.
- [8] H.Th. Jongen, D. Klatté, and K. Tammer. Implicit functions and sensitivity of stationary points. *Mathematical Programming*, 49:123–138, 1990.
- [9] C. Kanzow, N. Yamashita, and M. Fukushima. New NCP-functions and their properties. *Journal of Optimization Theory and Applications*, 94:115–135, 1997.
- [10] A. Kaplan. On the convergence of the penalty function method. *Soviet Math. Dokl.*, 17:1008–1012, 1976.
- [11] A. Kaplan and R. Tichatschke. *Stable Methods for ill-posed Variational Problems*. Akademie Verlag, Berlin, 1994.
- [12] A. King and R.T. Rockafellar. Sensitivity analysis for nonsmooth generalized equations. *Mathematical Programming*, 55:341–364, 1992.
- [13] D. Klatté and B. Kummer. Generalized Kojima functions and Lipschitz stability of critical points. *Computational Optimization and Applications*, 13:61–85, 1999.
- [14] D. Klatté and B. Kummer. *Nonsmooth Equations in Optimization - Regularity, Calculus, Methods and Applications*. Nonconvex Optimization and Its Applications. Kluwer Academic Publ., Dordrecht-Boston-London, 2002.

- [15] M. Kojima. Strongly stable stationary solutions in nonlinear programs. In S.M. Robinson, editor, *Analysis and Computation of Fixed Points*, pages 93–138. Academic Press, New York, 1980.
- [16] B. Kummer. Lipschitzian inverse functions, directional derivatives and application in $C^{1,1}$ optimization. *Journal of Optimization Theory and Applications*, 70:559–580, 1991.
- [17] B. Kummer. Newton’s method based on generalized derivatives for nonsmooth functions: convergence analysis. In W. Oettli and D. Pallaschke, editors, *Advances in Optimization*, pages 171–194. Springer, Berlin, 1992.
- [18] B. Kummer. Lipschitzian and pseudo-Lipschitzian inverse functions and applications to nonlinear programming. In A.V. Fiacco, editor, *Mathematical Programming with Data Perturbations*, pages 201–222. Marcel Dekker, New York, 1997.
- [19] W. Li. Abadie’s constraint qualification, metric regularity, and error bounds for differentiable convex inequalities. *SIAM Journal on Optimization*, 7:966–978, 1997.
- [20] C. Großmann und H. Kleinmichel. *Verfahren der nichtlinearen Optimierung*. Teubner-Texte zur Mathematik. Teubner Verlagsgesellschaft, Leipzig, 1976.
- [21] C. Großmann und J. Terno. *Numerik der Optimierung*. Teubner Studienbücher. Teubner Verlagsgesellschaft, Stuttgart, 1993.
- [22] R. Mifflin. Semismooth and semiconvex functions in constrained optimization. *SIAM Journal on Control and Optimization*, 15:957–972, 1977.
- [23] B.S. Mordukhovich and Y. Shao. Mixed coderivatives of set-valued mappings in variational analysis. *Journal of Applied Analysis*, 4:269–294, 1998.
- [24] J.-S. Pang and L. Qi. Nonsmooth equations: motivation and algorithms. *SIAM Journal on Optimization*, 3:443–465, 1993.
- [25] L. Qi and J. Sun. A nonsmooth version of Newton’s method. *Mathematical Programming*, 58:353–367, 1993.
- [26] S.M. Robinson. Strongly regular generalized equations. *Mathematics of Operations Research*, 5:43–62, 1980.
- [27] R.T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, N.J., 1970.
- [28] R.T. Rockafellar and R. J.-B. Wets. *Variational Analysis*. Springer, Berlin, 1998.
- [29] J. B. Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica*, 33:520–534, 1965.
- [30] D. Sun and L. Qi. On NCP functions. *Computational Optimization and Appl.*, 13:201–220, 1999.