

Verallgemeinerte Ableitungen in der nichtglatten Analysis: Zusammenhänge und Berechenbarkeit

Diplomarbeit

Humboldt-Universität zu Berlin
Mathematisch-Naturwissenschaftliche Fakultät II
Institut für Mathematik

23.März 2009

eingereicht von: Mario Krell
geboren am: 25.02.1984 in: Frankfurt (Oder)
Betreuer: Prof. Dr. Kummer

Inhaltsverzeichnis

1	Vorwort	5
2	Definitionen und Verallgemeinerungen	7
2.1	Kegelkonzepte	7
2.2	Verallgemeinerte Ableitungen	10
2.3	Klassische Ableitungsbegriffe	12
2.4	Was wird wodurch verallgemeinert?	13
3	Beispiele	13
4	Allgemeine Sätze	19
4.1	Ekeland's Variationsprinzip	19
4.2	Konvexe Analysis	20
4.3	Fréchet-Konzept	21
4.4	Clarke-Konzept	23
5	Teilmengenbeziehungen	24
5.1	Vergleich der Normalenkegelkonzepte	24
5.2	Beziehungen zwischen Normalenkegel und Subdifferential	26
5.3	Beziehungen bzgl. Co-Differential, Kontingent- und Thibault-Ableitung	28
6	Erste Rechenregeln	32
6.1	...für Normalenkegel	32
6.2	...für die Co-Differentiale	33
6.3	...für die Subdifferentiale	34
6.4	...für die verallgemeinerte Ableitung der Umkehrabbildung	37
6.5	...für Thibault- und Kontingent-Ableitung	37
7	Stabilität - Co-Differential - Regularität	43
7.1	Aubin-Eigenschaft	44
7.2	Limiting Epsilon-Normalen und das Co-Differential	46
7.3	Mordukhovich-Kriterium	47
7.4	Limiting-Fréchet-Regularitätsbedingung bzw. GMFCQ	49
8	Anwendung der GMFCQ - Hauptrechenregel	50
8.1	Vergleich der Anwendung der GMFCQ mit der klassischen MFCQ	53
9	Optimalitätsbedingungen	55
9.1	„Gegenbeispiele“ für Optimalitätsbedingungen	57
10	Konkrete Berechnungen	58
10.1	...bei Differenzierbarkeit	58
10.2	...bei „Regularität“	59

10.2.1	Normalenkegel	59
10.2.2	Optimalität, Lagrange-Funktion, Karush-Kuhn-Tucker-Bedingung	60
10.2.3	Die Ableitung der Kojima-Funktion	61
10.2.4	Die Ableitung der Funktion der stationären Punkte	63
10.3	...mit Maximumtypfunktionen	64
10.3.1	Definition und Subdifferential	64
10.3.2	Die Kontingent-Ableitung des Subdifferentials	65
10.4	...mit stückweise differenzierbaren Funktionen	66
10.4.1	Definitionen und Stetigkeit	66
10.4.2	Subdifferential und Richtungsableitung	68
10.4.3	Weitere Ableitungen	71
10.4.4	Anwendung auf Normalenkegel und Optimalität	72
10.5	...zur Normalenkegelabbildung zum positiven Orthanten	78
11	Mehrstufenoptimierung	78
11.1	GMPEC	78
11.2	GMPCC	79
11.3	Vergleich von GMPEC und GMPCC	80
12	Formelsammlung	82
13	Thesen	87
14	Symbolverzeichnis	89
15	Literatur	90
16	Erklärungen	91

1 Vorwort

Ich möchte mit einer Betrachtung der Optimierungsaufgaben aus der Schule beginnen. Hier lernt man, dass die potentiellen Extrempunkte einer Funktion sich an den Stellen befinden, wo die Ableitung verschwindet. Ist eine Zulässigkeitsmenge für die Funktion gegeben, im folgenden meist M genannt, so muss man die Zulässigkeit optimalverdächtiger Punkte prüfen. Außerdem muss man untersuchen, ob die Randwerte der Zulässigkeitsmenge, die meist ein abgeschlossenes Intervall ist, Minima oder Maxima sein können. Der ein oder andere lernt sogar schon mit dem Gradienten bzw. mit vektorwertigen Funktionen umzugehen. Im Studium prallen dann mehrere Themengebiete aufeinander. Hört man etwas aus der Linearen Optimierung, so erfährt man, dass man die Ecken der Restriktionsmenge untersuchen muss und lernt das Simplexverfahren kennen. Weitere Themengebiete sind Mannigfaltigkeiten, zugehörige Differenzierbarkeitsstrukturen, Tangentialräume und Normalenvektorfelder. Letztendlich lernt man womöglich noch verschiedene Ableitungsbegriffe kennen und was Konvexität für eine tolle Eigenschaft ist, weil man so viel daraus ableiten kann und weil bei konvexen Funktionen jedes lokale Minimum ein globales ist.

In dieser Arbeit geht es um das, was „danach“ kommt. Die Restriktionsmenge kann eine eher unschöne Gestalt und eine kompliziertere Beschreibung haben, Tangentialräume sind nicht mehr schön ermittelbar, Konvexität und lokale Minima, die gleichzeitig globale Minima sind, findet man nur noch selten und Funktionen sind nicht mehr normal differenzierbar. Aber weiterhin will man Optimierungsaufgaben lösen, das heißt, auf einer Restriktionsmenge M das Minimum einer gegebenen Funktion bestimmen. Hierbei kann erschwerend hinzukommen, dass die Restriktionsmenge unter anderem aus lokalen oder globalen Minimierern bzw. optimalverdächtigen Punkten einer anderen Optimierungsaufgabe besteht.

Um diese Aufgaben nun zu lösen, gibt es zahlreiche Ansätze, die oft auf bestimmte günstige Voraussetzungen oder Eigenschaften zugeschnitten sind. Die aus meiner Sicht wichtigsten Ansätze werde ich in meiner Arbeit zunächst vorstellen und einen Überblick darüber geben, wie sie zusammenhängen. Dabei verwende ich im wesentlichen die Werke von F. Clarke, R.T. Rockafellar und B. Kummer.

Auch wenn viele Aussagen und Definitionen in beliebigen Räumen existieren, beschränke ich mich auf den endlichdimensionalen Fall $X = \mathbb{R}^n$, da dieser allein schon recht schwierig zu behandeln ist. Will man konkrete Berechnungen anstellen, was ein weiteres meiner Ziele ist, so stößt man außerdem im unendlichdimensionalen Fall recht schnell an seine Grenzen.

Um mich dann mit konkreten Sätzen auseinander zu setzen, mit komplexeren Restriktionsmengen und Berechnungen, werde ich einige Ansätze aus einem Artikel von J. Outrata [Out06] betrachten und mit anderen Konzepten vergleichen, wie sie zum Beispiel S. Scholtes [SS00] benutzt. Außerdem werde ich untersuchen, was diese Aussagen für die konkrete Beschreibung von Mengen und Funktionen bzw. für konkreter beschriebene Mengen und Funktionen bedeuten.

Wie meist üblich, werde ich nur abgeschlossene Restriktionsmengen M betrachten. Alles andere wäre nicht nur zu kompliziert, sondern ist auch zumeist nur schwer beschreibbar

und im Allgemeinen für die Praxis nicht so relevant. Das Hauptproblem aus meiner Sicht ist jedoch, dass man sich dem lokalen Minimum meist über umgebende Punkte nähern können will. Dies ist aber nur bei abgeschlossenen Mengen möglich. Außerdem weiß man von stetigen Funktionen, dass sie auf beschränkten, abgeschlossenen Mengen im \mathbb{R}^n sowohl ein globales Minimum als auch ein globales Maximum annehmen.

Bei der Strukturierung meiner Arbeit möchte ich zunächst die Grundlagen klären bzw. wiederholen. Dann möchte ich immer komplexere Zusammenhänge darstellen. Das heißt, zunächst zeige ich einfache Beispiele, dann betrachte ich einfache Zusammenhänge und danach komplexere Zusammenhänge zwischen den Ableitungen. Als nächste Herausforderung werde ich konkrete Formeln suchen und dazu zunehmend komplexere Abbildungen betrachten. Letztendlich erfolgt dann eine kleine Auseinandersetzung mit mehrstufiger Optimierung.

Danksagung

Da ich verpflichtet bin, alle „Hilfsmittel“, die ich „verwendet“ habe, aufzuzählen, möchte ich hier die Gelegenheit nutzen um mich bei allen zu bedanken, die direkt oder indirekt bei der Entstehung dieser Arbeit geholfen haben.

Zuerst sind meine Eltern zu nennen. Sie haben mir die nötigen Mathematikgene mitgegeben und mich immer mit vollem Einsatz unterstützt. Frau Böhm, Herr Bochmann und Herr Mettke waren außergewöhnliche Lehrer, die meine Fähigkeiten für die Mathematik sehr stark geprägt haben. Wichtig für die Entscheidung ein Mathematikstudium zu beginnen, waren außer meinen Eltern und Herrn Mettke noch Volker Grabsch, aber vor allem Thomas Körner. Ich habe diese Entscheidung nie bereut und bin froh, dass es dazu kam.

Im Studium haben vor allem Professor Friedrich in der Analysis und Professor Kummer in der Optimierung die fachlichen Grundlagen für die Diplomarbeit gelegt. Während des Studiums hatte ich die Unterstützung von vielen Freunden. Insbesondere Volker Grabsch hat mich unterstützt. Eine Unterstützung im Studium war auch die Hans-Böckler-Stiftung, speziell Birgit Grafe. Sie hat mich immer sehr gut beraten in der Organisation der Endphase meines Studiums. Immer ein offenes Ohr und gute Hinweise für mich hatten Jan Heerda, Paul Ebermann, Christian Fleck, Thomas Ueckerdt und Clarissa Meyer.

Alle diesen Personen danke ich für ihre Unterstützung. Ich danke auch Professor Griewank, dass er sich ohne zu zögern bereit erklärt hat, das zweite Gutachten für diese Arbeit zu schreiben.

Ganz besonderer Dank gilt Professor Kummer. Er hat mir nicht nur den schönsten Bereich der Mathematik nahegebracht und mir ein Diplomthema gegeben, wie ich mir kein besseres hätte vorstellen können, sondern er hat mich sehr gut betreut mit zahlreichen Anregungen für den Inhalt der Diplomarbeit und vielen Antworten auf fachliche Fragen. Ohne diese Unterstützung wäre die Arbeit wohl nicht so schnell fertig geworden und hätte nicht diesen großen Umfang.

2 Definitionen und Verallgemeinerungen

In diesem Abschnitt werde ich alle für meine Arbeit relevanten Definitionen verallgemeinerter Ableitungen einführen. Dies dient dazu, dem mit der nichtglatten Analysis weniger vertrauten Leser die grundlegenden Begriffe nahe zu bringen. Der erfahrene Leser kann sich damit genügen, in diesem Abschnitt die Notation nachzuschlagen, da diese in diesem Fachgebiet leider nicht einheitlich ist. Es steht außerdem am Ende der Arbeit ein Symbolverzeichnis zur Verfügung.

Ich habe sehr darauf Wert gelegt, in meiner Notation keine Konzepte abzuwerten, ähnliche Konzepte ähnlich zu bezeichnen und die Notation abhängig von einer Bezeichnung zu wählen.

Da der zugrundeliegende Vektorraum X der \mathbb{R}^n ist, ist der zugehörige Dualraum wieder X . Generell sei f eine Funktion mit $f : X \rightarrow \bar{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$. Die Möglichkeit negativ-unendlicher Werte lasse ich bewusst weg und auch der Funktionswert „ ∞ “ tritt im Allgemeinen nicht auf bzw. ich betrachte nur die Menge, auf der f endlich ist. M sei die Restriktionsmenge auf die wir uns beziehen und $x \in M$ das Element in dem wir die definierte Menge betrachten. Solange nicht explizit anders erwähnt, gehen ich davon aus, dass die Restriktionsmenge abgeschlossen ist.

2.1 Kegelkonzepte

In der nichtglatten Optimierung gibt es sehr viele Konzepte für Tangenten und Normalen. Die Bekanntesten werden im folgenden aufgezählt. Dazu werden nicht Untervektorräume verwendet sondern Kegel, d.h. Mengen $A \subset X$ mit $A = \mathbb{R}_+ A$. Eine wichtige Rolle spielt hierbei der Polarkegel.

Definition 2.1 (Polarkegel) Der Polarkegel zur Menge $A \subset X$ ist der Kegel

$$A^* = \{v \in X \mid \langle v, u \rangle \leq 0 \ \forall u \in A\}.$$

Der Polarkegel verallgemeinert die Eigenschaft, dass der Normalenraum senkrecht zum Tangentialraum steht. Wäre der Tangentialraum ein Untervektorraum, so könnte man „ $=$ “ statt „ \leq “ schreiben und hätte die klassische Orthogonalitätsbedingung, welche nun durch eine „Polarbedingung“ verallgemeinert wird.

Zuvor ist natürlich noch zu klären, was die Kegel mit verallgemeinerten Ableitungen zu tun haben, da dieser Zusammenhang nicht sofort ersichtlich ist.

Zum einen benötigt man die Kegelkonzepte für einige Definitionen der verallgemeinerten Ableitungen. Zum zweiten sind die Kegel wichtig, um besser mit Funktionen umgehen zu können, die einen eingeschränkten Definitionsbereich haben bzw. bei denen „ ∞ “ als Funktionswert zugelassen ist. Und zum dritten werden wir später sehen, dass die Normalenkegel verallgemeinerte Ableitungen (Subdifferentialle) von speziellen Funktionen (Indikatorfunktionen) sind, was in starken Zusammenhang zum zweiten Punkt steht.

Der ursprüngliche Normalenkegel stammt aus der konvexen Optimierung.

Definition 2.2 (Kegel der konvexen Optimierung) Der Normalenkegel an die Menge M im Punkt x ist

$$N_M^{ko}(x) = (M - x)^*.$$

Möglichkeiten für den zugehörigen Tangentialkegel wären

$$T_M^{ko}(x) = M - x, \mathbb{R}_+(M - x), \overline{\mathbb{R}_+(M - x)} \text{ oder } \overline{\mathbb{R}_+(\text{conv}(M - x))},$$

wobei der resultierende polare Normalenkegel immer der selbe wäre.

Definition 2.3 (Mengenprodukt) Seien $A, B \subset X$ zwei beliebige Mengen. Dann bezeichnet

$$AB = \{x \mid x = a \cdot b, a \in A, b \in B\}$$

das elementweise Produkt der beiden Mengen.

Später werden wir sehen, dass von allen folgenden Normalenkegeln dieser Normalenkegel der kleinste ist und $\overline{\mathbb{R}_+(\text{conv}(M - x))}$ der größte Tangentialkegel ist. Ziemlich naheliegende Verallgemeinerungen sind die folgenden zwei Normalenkegel, welche sehr stark miteinander verwandt sind.

Definition 2.4 (Kontingenz-Kegel)

$$\begin{aligned} T_M^B(x) &= \left\{ u \in X \mid \liminf_{t \searrow 0} \frac{\text{dist}(x + tu, M)}{t} = 0 \right\} \\ &= \{ u \in X \mid \exists u_k \rightarrow u, t_k \searrow 0 : x + t_k u_k \in M \} \end{aligned}$$

$t \searrow 0$ entspricht dabei $t \rightarrow 0$ und $t > 0$. In der ganzen Arbeit gilt wie üblich $t \in \mathbb{R}$.

$N_M^B(x) = T_M^B(x)^*$ ist der Kontingenz-Normalenkegel.

Oftmals wird statt $T_M^B(x)$ auch die Bezeichnung $C_M(x)$ verwendet, wahrscheinlich abgeleitet von „contingent cone“, oder auch „tangent cone“ mit der Notation $T_M(x)$. Ein weiterer Name ist Bouligand-Kegel, woher die gewählte Notation stammt.

Definition 2.5 (Fréchet-Normalenkegel)

$$N_M^F(x) = \{ v \mid \langle v, y - x \rangle \leq o(y - x) \ \forall y \in M \}$$

Dieser Kegel wird auch „F-normal cone“ genannt [Kum08] oder „regular normal cone“ und mit $\hat{N}_M(x)$ bezeichnet [RW98].

Im Vergleich zum konvexen Normalenkegel interessiert die Menge M nur noch in einer kleinen Umgebung von x . Zudem wird nicht mehr $\langle v, y - x \rangle \leq 0$ gefordert, sondern eine kleine „Störung“ ist zulässig.

Im folgenden werden dieser Normalenkegel in einer kleinen Umgebung und die „Grenzwerte“ der Normalenkegel betrachtet.

Definition 2.6 (Limiting-Fréchet-Normalenkegel)

$$N_M^{\text{lim},F}(x) = \left\{ x^* \in X \mid \exists x_k \xrightarrow{M} x, x_k^* \rightarrow x^* : x_k^* \in N_M^F(x_k) \right\}$$

Mitunter schreibt man auch

$$N_M^{lim,F}(x) = \limsup_{x_k \xrightarrow{M} x} N_M^F(x_k),$$

wobei hier der äußere Grenzwert (\limsup) für den Grenzwert mengenwertiger Abbildungen im Sinne von Hausdorff bzw. Painléve-Kuratowski steht. Weitere auftretende Bezeichnungen sind „limiting F-normal cone“, „limiting normal cone“ oder nur „normal cone“ mit der Notation $N_M(x)$.

Definition 2.7 (normal regulär) M wird als in x normal regulär bezeichnet, wenn $N_M^{lim,F}(x) = N_M^F(x)$ gilt, also Fréchet- und Limiting-Fréchet-Normalenkegel gleich sind.

Definition 2.8 (abgeschlossen) Eine mengenwertige Abbildung $F : X_1 \rightrightarrows X_2$ wird als abgeschlossen bezeichnet, wenn ihr Graph,

$$\text{Gph } F := \{(x, y) \in X_1 \times X_2 \mid y \in F(x)\},$$

abgeschlossen ist im Sinne der Kreuzmetrik, das heißt, wir addieren komponentenweise die Abstände. Manchmal nimmt man auch die Wurzel aus der Summe der Quadrate der beiden Abstandskomponenten. Da wir im \mathbb{R}^n sind, ist egal welche Variante man wählt oder ob man gleich eine Norm auf dem gesamten Raum nimmt. Schließlich sind alle Normen im \mathbb{R}^n äquivalent und somit auch die entsprechenden induzierten Topologien.

Abgeschlossenheit ist eine Verallgemeinerung der Oberhalbstetigkeit für einwertige Abbildungen. Es gibt noch einige weitere Verallgemeinerungen, welche aber für unsere Betrachtungen nicht relevant sind.

Eine Diagonalisierungsargumentation zeigt, dass $N_M^{lim,F}(x)$ abgeschlossen ist bzgl. x . „Stetigkeit bzgl. M “ oder „Unterhalbstetigkeit“ gilt im allgemeinen nicht.

Betrachtet man die Normalenkegel als Abbildungen auf M , so ist $N_M^{lim,F}(x)$ als die kleinste abgeschlossene Abbildung zu sehen, die $N_M^F(x)$ bzw. die zugehörigen Bilder enthält. Interessanterweise ändern sich durch den Grenzwertübergang viele Eigenschaften bzw. kehren sich um, wie später immer wieder zu sehen sein wird.

Zum bereits definierten Kontingent-Kegel kann man eine andere Art Grenzwertübergang machen und erhält einen neuen, kleineren Kegel.

Definition 2.9 (Clarke-Tangentialkegel)

$$T_M^{Cl}(x) = \left\{ u \in X \mid \forall x_k \xrightarrow{M} x, t_k \searrow 0, \exists u_k \rightarrow u : x_k + t_k u_k \in M \right\}$$

Der Clarke-Normalenkegel $N_M^{Cl}(x)$ ist der zugehörige Polarkegel.

Später werden wir sehen, dass $N_M^{Cl}(x)$ die kleinste konvexwertige abgeschlossene Abbildung ist, die $N_M^F(x)$ bzw. die zugehörigen Bilder enthält. Durch die Zusammenhänge der Normalenkegel, die wir später zeigen werden, folgt, dass man in der Definition von „normal regulär“ $N_M^{Cl}(x)$ und $N_M^{lim,F}(x)$ beliebig austauschen kann.

Die ursprüngliche Definition des Tangentialkegels, zu finden in [Cla83] S.51, war etwas anders. Man brauchte dazu die Clarke-Richtungsableitung im betrachteten Punkt von der Funktion, die den Abstand eines Punktes zur Menge M angab.

Definition 2.10 (Clarke-Richtungsableitung)

$$f^{Cl}(x; u) = \limsup_{x_k \rightarrow x, t_k \searrow 0} \frac{f(x_k + t_k u) - f(x_k)}{t_k}$$

Der Tangentialraum ergab sich dann aus den Richtungen mit Ableitung Null zur Abstandsfunktion betrachtet im Punkt x . Die von mir verwendete Version ist wesentlich einfacher, gebräuchlicher und äquivalent zu der ursprünglichen.

2.2 Verallgemeinerte Ableitungen

Eine von vielen verallgemeinerten Ableitungen ist die eben erwähnte Clarke-Richtungsableitung. Sie kann verwendet werden um das zugehörige Subdifferential zu definieren. Hier gehen wir allerdings den Weg über die Normalenkegel, welcher die Zusammenhänge der Konzepte besser zeigt.

Der Begriff des Subdifferentials wurde durch die konvexe Optimierung geprägt, wo in einem Punkt x_0 die Subgradienten ($g \in \partial f$) die linearen Tangentialunterräume

$$\{g(x - x_0) + f(x_0) | x \in X\}$$

bilden, die man an den Epigraphen im Punkte x_0 legen kann.

Definition 2.11 (Subdifferential zum Normalenkegel N^α)

$$\partial^\alpha f(x) = \left\{ x^* \mid (x^*, -1) \in N_{\text{epi } f}^\alpha(x, f(x)) \right\}$$

Zur Erinnerung sei nochmal wiederholt,

Definition 2.12 (Epigraph)

$$\text{epi } f = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq t\}.$$

Analog zum Normalenkegel definiert man Regularität für Funktionen.

Definition 2.13 (reguläre Funktion) Eine beliebige Funktion f wird regulär in x genannt, wenn die drei verallgemeinerten Subdifferentialie in x übereinstimmen. Gilt dies in jeden Punkt, so sagt man f ist regulär.

Bemerkung Später werden wir sehen, dass dies der Fall ist, wenn das Subdifferential $\partial^F f(x)$ mit $\partial^{\text{lim}, F} f(x)$ oder $\partial^{Cl} f$ übereinstimmt. Am Beispiel der Sägezahnfunktion sehen wir, dass die Betrachtung von $\partial^F f$ notwendig ist. Auch werden wir beweisen, dass konvexe, stetig differenzierbare und Funktionen vom Maximumtyp regulär sind. In Abschnitt 3 gibt es einige Beispiele nicht regulärer Funktionen.

Ist das Subdifferential nur für reelle Funktionale definiert, so sind die vier folgenden verallgemeinerten Ableitungen auch für Abbildungen in den \mathbb{R}^m gedacht. Sie werden für verschiedene Stabilitätsaussagen benötigt. Für die folgenden drei kann die Abbildung $F : X \rightrightarrows Y := \mathbb{R}^m$ sogar mehrwertig sein.

Definition 2.14 (Co-Differential)

$$D^*F(x, y)(y^*) = \left\{ x^* \in X \mid (x^*, -y^*) \in N_{\text{Gph } F}^{\text{lim}, F}(x, y) \right\}$$

Für die Co-Differentiale bzgl. anderer Normalenkegel benutzt ich die Bezeichnung $D^\alpha F(x, y)(y^*)$, mit $\alpha \in \{F; Cl\}$. Dabei werde ich im Vergleich dann auch „lim,F“ verwenden. $D^F F(x, y)(y^*)$ wird als $\hat{D}^*F(x, y)(y^*)$ in [RW98] untersucht. $D^{Cl}F(x, y)(y^*)$ tritt gar nicht auf, da es im Allgemeinen zu groß ist und keine brauchbaren Aussagen liefert.

Definition 2.15 (Kontingenz-Ableitung)

$$CF(x, y; u) = \{v \in X \mid \exists t_k \searrow 0, (u_k, v_k) \rightarrow (u, v) : y + t_k v_k \in F(x + t_k u_k)\}$$

Ist F einwertig, so definiert man auch analog

$$CF(x; u) = \left\{ v \in X \mid \exists t_k \searrow 0, u_k \rightarrow u : v = \lim_{k \rightarrow \infty} \frac{F(x + t_k u_k) - F(x)}{t_k} \right\}.$$

Diese Ableitung wird auch Bouligand- oder graphische Ableitung (Notation: $DF(x|y)(u)$ [RW98]) genannt. Es gilt nach Definition:

$$v \in CF(x, y; u) \Leftrightarrow (u, v) \in T_{\text{Gph } F}^B(x, y).$$

Somit könnte man auch analog eine Ableitung über den Clarke-Tangentialekegel definieren. Alternativ kann man natürlich wieder einen Grenzwertansatz betrachten, wie den folgenden:

Definition 2.16 (Thibault-Ableitung) $TF(x, y; u)$

$$= \left\{ v \in X \mid \exists (x_k, y_k) \xrightarrow{\text{Gph } F} (x, y), \exists t_k \searrow 0, (u_k, v_k) \rightarrow (u, v) : y_k + t_k v_k \in F(x_k + t_k u_k) \right\}$$

Die einwertige Variante lautet diesmal

$$TF(x; u) = \left\{ v \in X \mid \exists x_k \rightarrow x, \exists t_k \searrow 0, u_k \rightarrow u : v = \lim_{k \rightarrow \infty} \frac{F(x_k + t_k u_k) - F(x_k)}{t_k} \right\}.$$

Diese Ableitung wird auch strikte graphische Ableitung genannt [RW98]. Der zugehörige Tangentialekegel wäre

$$T_M^T(x) = \left\{ u \mid \exists x_k \xrightarrow{M} x, t_k \searrow 0, u_k \rightarrow u : x_k + t_k u_k \in M \right\}.$$

Der Unterschied zum Clarke-Tangentialekegel besteht in den anderen Quantoren zu den t_k und x_k . Dieser Kegel ist im Allgemeinen sogar größer als der Kontingenzkegel. Mir ist nicht bekannt, dass der zugehörige Normalenkegel betrachtet wird.

Bemerkung (Lipschitzfall) In beiden Ableitungen werden wir für lokal lipschitzstetige Funktionen die $u_k \equiv u$ setzen. Das dies möglich ist, zeigt eine einfache Abschätzung. Sei L eine Lipschitzkonstante um x . Dann gilt:

$$\|(f(x_k + t_k u_k) - f(x_k)) - (f(x_k + t_k u) - f(x_k))\| \leq L t_k \|u_k - u\|.$$

Außerdem werden viele Beweise mit x_k geführt und den Satz mit der Kontingent-Ableitung erhält man mittels $x_k \equiv x$.

Der bekannte Satz von Rademacher besagt, dass eine lokal lipschitzstetige Abbildung von einer offenen Menge in den \mathbb{R}^m fast überall (im Sinne des Lebesguemaßes) total differenzierbar ist. Diesen Effekt macht sich F.H. Clarke zunutze für einen Grenzwertübergang. Sei $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ nun wieder einwertig.

Definition 2.17 (verallgemeinerte Jakobimatrix) Diese wird auch verallgemeinerte Clarke-Ableitung genannt.

$$\partial^{Cl}(F)(x) = \text{conv} \{v = \lim DF(x_k) | x_k \rightarrow x, DF(x_k) \text{ existiert}\}$$

2.3 Klassische Ableitungsbegriffe

Es folgen nun zum Vergleich die klassischen Ableitungsbegriffe.

Während man bei den vorhergehenden Definitionen im schlimmsten Fall leere Mengen erhält, ist hier zu beachten, dass die Ableitungen eventuell sogar nicht definiert sind. Außerdem gelten sie nicht für mehrwertige Abbildungen. Gelte also $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Definition 2.18 (Ableitung) Diese wird auch „totales Differential“ genannt. In Banachräumen verwendet man die Bezeichnung „Fréchet-Ableitung“.

$$DF(x) \in L(\mathbb{R}^n, \mathbb{R}^m) : \lim_{\|h\| \searrow 0} \frac{\|F(x+h) - F(x) - DF(x)h\|}{\|h\|} = 0$$

Mitunter wird für DF auch die Notation ∇F verwendet. Betrachte ich die Ableitung nur bezüglich einer bestimmten Variable der Funktion, so kennzeichne ich dies speziell, indem ich die Variable als Index nach dem D einfüge wie zum Beispiel bei $D_x F(x, y)$. Wird nur nach der i -ten Komponente von F abgeleitet, so schreibe ich DF_i . DF^T bezeichnet die Duale Abbildung bzw. die transponierte Matrix.

Da X endlichdimensional ist, identifiziere ich im Falle $m = 1$ „ DF “ manchmal mit seinem Gradienten, ähnlich wie ich bei den Normalenkegeln den Vektorraum statt dem Dualraum benutze. Die jeweilige Verwendung ergibt sich dann immer aus dem Kontext.

Definition 2.19 (Gâteaux-Differential)

$$dF(x; u) = \lim_{t \rightarrow 0} \frac{F(x + tu) - F(x)}{t}$$

Ist hier u der Einheitsvektor e_i , so nennt man die Ableitung auch partielle Ableitung und bezeichnet sie mit $\frac{\partial F}{\partial x_i}$.

Definition 2.20 (Richtungsableitung) Auch einseitiges Gâteaux-Differential genannt.

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

2.4 Was wird wodurch verallgemeinert?

- Das Subdifferential ist sowohl eine Verallgemeinerung des Subdifferentials der konvexen Analysis als auch des allgemeinen Differentials.
- Die Tangentialräume bzw. -mengen sind daher interessant, dass sie in der Differentialgeometrie die Richtungen angeben, in denen eine Funktion abgeleitet bzgl. einer Mannigfaltigkeit wird, mit der ja unsere Restriktionsmengen in gewisser Weise vergleichbar sind.
- Die Orthogonalität des Normalenraums zum Tangentialraum wird durch den Polarkegel verallgemeinert.
- Die Normalen- und Tangentialkegelkonzepte sind natürlich aus der Sicht relevant, dass sie die Grundlage für die anderen Ableitungskonzepte bilden und für wichtige Sätze benötigt werden.
- Das Co-Differential kann man als Verallgemeinerung des Subdifferentials ansehen, da wir $y^* = 1$ als Richtung wählen können. Außerdem kann man zu einer Funktion f eine mehrwertige Abbildung F definieren mit $F(x) = \{(x, t) | f(x) \leq t\}$. Dann gilt $D^\alpha F(x, f(x))(1) = \partial^\alpha f(x)$.
- Kontingent- und Thibault-Ableitung sind Verallgemeinerungen der Richtungsableitung. Hierbei kann man die Kontingent-Ableitung dem Bouligand-Konzept zuordnen.
- Die Richtungsableitung kann man wiederum der konvexen Optimierung zuordnen bzw. speziell für den konvexen Fall betrachten. Man kann sie auch als Verallgemeinerung der anderen klassischen Ableitungsbegriffe sehen und umgekehrt.
- Die Clarke-Richtungsableitung ist letztendlich das Supremum der Thibault-Ableitung.

Der Vollständigkeit halber muss erwähnt werden, dass dies nicht alle Konzepte sind. Es gibt noch viele andere Verallgemeinerungen und natürlich betrachtet man die genannten Konzepte auch auf allgemeineren Räumen. Allerdings wird dann die Berechnung im Allgemeinen noch wesentlich schwieriger. In dieser Arbeit beschränke ich mich auf die genannten Konzepte, da sie aus meiner Sicht die wichtigsten sind.

3 Beispiele

Um ein besseres Gefühl für die Begriffe und diesbezügliche Fragestellungen zu bekommen, ist es immer wichtig konkrete Beispiele zu betrachten. Beispiele haben außerdem den Vorteil, dass man erkennt, was für Informationen die zahlreichen Ableitungen enthalten und wie sie auch zum Teil nichtssagend sein können. Ich habe die Ableitungen nur in dem Punkt betrachtet, wo sie interessant werden und Unterschiede aufzeigen.

Das erste Beispiel zeigt, wie bei einer konvexen Funktion die Subdifferenziale zusammenfallen, aber sich bei den verallgemeinerten Ableitungen die Ergebnisse unterscheiden.

Beispiel 1 ($f(x) = |x|$) ($\alpha \in \{ko; F; lim, F; Cl\}$)

$$\begin{aligned}
N_{\text{epi } f}^{\alpha}(0, 0) &= \{(u, v) | v \leq -|u|\} \\
\partial^{\alpha} f(0) &= [-1, 1] \\
D^{Cl} f(0, 0)(v) = \mathbb{R}, \quad D^F f(0, 0)(v) &= \{u | -v \leq -|u|\} \\
D^{lim, F} f(0, 0)(v) &= \{u | -v \leq -|u| \text{ oder } -v = |u|\} \\
Cf(0, 0; u) &= \{|u|\} \\
Tf(0, 0; u) &= [-|u|, |u|]
\end{aligned}$$

Das folgende Beispiel zeigt, wie sich durch einen Vorzeichenwechsel nun auch die Subdifferenziale unterscheiden.

Beispiel 2 ($f(x) = -|x|$)

$$\begin{aligned}
N_{\text{epi } f}^{ko}(0, 0) = N_{\text{epi } f}^F(0, 0) &= \{0\} \\
N_{\text{epi } f}^{lim, F}(0, 0) &= \{(u, -|u|) | u \in \mathbb{R}\} \\
N_{\text{epi } f}^{Cl}(0, 0) &= \{(u, v) | v \leq -|u|\} \\
\partial^{ko} f(0) = \partial^F f(0) &= \emptyset \\
\partial^{lim, F} f(0) &= \{-1, 1\} \\
\partial^{Cl} f(0) &= [-1, 1] \\
D^{Cl} f(0, 0)(v) = \mathbb{R}, \quad D^F f(0, 0)(v) &= \{u | -v \geq |u|\} \\
D^{lim, F} f(0, 0)(v) &= \{u | -v \geq |u| \text{ oder } -v = -|u|\} \\
Cf(0, 0; u) &= \{-|u|\} \\
Tf(0, 0; u) &= [-|u|, |u|]
\end{aligned}$$

Am nächsten Beispiel sieht man, dass bei stetig differenzierbaren Funktionen die Ableitungen zusammenfallen, bis auf das konvexe Subdifferential, welches dann sogar leer sein kann.

Beispiel 3 ($f(x) = -x^2$)

$$\begin{aligned}
N_{\text{epi } f}^{ko}(0, 0) &= \{0\} \\
N_{\text{epi } f}^F(0, 0) = N_{\text{epi } f}^{lim, F}(0, 0) = N_{\text{epi } f}^{Cl}(0, 0) &= \{0\} \times \mathbb{R}_- \\
\partial^{ko} f(0) &= \emptyset \\
\partial^F f(0) = \partial^{Cl} f(0) = \partial^{lim, F} f(0) &= \{0\} \\
D^F f(0, 0)(v) = D^{lim, F} f(0, 0)(v) = D^{Cl} f(0, 0)(v) &= \{0\} \\
Cf(0, 0; u) = Tf(0, 0; u) &= \{0\} \\
Df(0)u &= 0
\end{aligned}$$

Hat man die entsprechenden Normalenkegel an den Graphen einer Funktion, so lassen sich viele verallgemeinerte Ableitungen ablesen. Da sich der Graph der Umkehrfunktion aus einer einfachen „Spiegelung“ ergibt, lassen sich die Normalenkegel und somit die entsprechenden Ableitungen einfach berechnen. Diesen Effekt nutzen wir nun für eine recht einfache differenzierbare Funktion.

Beispiel 4 ($f(x) = x^3, g(x) = \text{sign}(x)|x|^{-3}$)

Für f :

$$\begin{aligned}
 N_{\text{epi } f}^{ko}(0, 0) &= \{0\} \\
 N_{\text{epi } f}^F(0, 0) = N_{\text{epi } f}^{lim, F}(0, 0) = N_{\text{epi } f}^{Cl}(0, 0) &= \{0\} \times \mathbb{R}_- \\
 \partial^{ko} f(0) &= \emptyset \\
 \partial^F f(0) = \partial^{Cl} f(0) = \partial^{lim, F} f(0) &= \{0\} \\
 D^F f(0, 0)(v) = D^{lim, F} f(0, 0)(v) = D^{Cl} f(0, 0)(v) &= \{0\} \\
 Cf(0, 0; u) = Tf(0, 0; u) &= \{0\} \\
 Df(0) &= 0
 \end{aligned}$$

Für g :

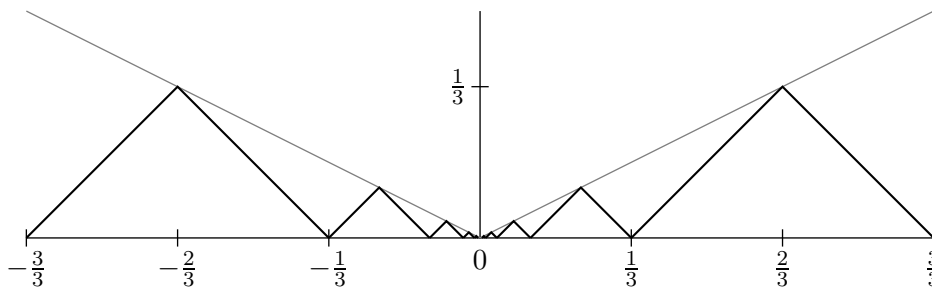
$$\begin{aligned}
 N_{\text{epi } g}^{ko}(0, 0) &= \{0\} \\
 N_{\text{epi } g}^F(0, 0) = N_{\text{epi } g}^{lim, F}(0, 0) = N_{\text{epi } g}^{Cl}(0, 0) &= \mathbb{R}_+ \times \{0\} \\
 \partial^{ko} g(0) = \partial^F g(0) = \partial^{Cl} g(0) = \partial^{lim, F} g(0) &= \emptyset \\
 D^F g(0, 0)(u) = D^{lim, F} g(0, 0)(u) = D^{Cl} g(0, 0)(u) &= \begin{cases} \mathbb{R}, & u = 0 \\ \emptyset, & u \neq 0 \end{cases} \\
 Cg(0, 0; v) = Tg(0, 0; v) &= \begin{cases} \mathbb{R}, & v = 0 \\ \emptyset, & v \neq 0 \end{cases}
 \end{aligned}$$

$Dg(0)$ existiert nicht.

Wie man schnell sieht, haben wir so ein Beispiel, wo alle Ableitungen entartet sind. Man kann somit auch sagen, dass sie nicht existieren, auch wenn das in diesem Zusammenhang unüblich ist.

Nun werden wir sehen, dass selbst bei „verrückten“ Funktionen f die Subdifferenziale $\partial^{lim, F} f(0)$ und $\partial^{Cl} f(0)$ zusammenfallen können, sich aber von $\partial^F f(0)$ unterscheiden. Viel interessanter bei den folgenden Beispielen ist jedoch der große Unterschied der Thibault- und der Kontingent-Richtungsableitung.

Beispiel 5 (Sägezahnfunktion $f(x)$ zu $g(x) = \frac{1}{2}|x|$) Sei $f : \mathbb{R} \rightarrow \mathbb{R}_+$ folgende stetige Funktion. In ihren differenzierbaren Punkten x_d habe sie den Anstieg 1 oder -1 und für die Funktionswerte gelte $0 \leq f(x_d) \leq g(x_d)$. Für die nichtdifferenzierbaren Punkte x_n gelte $f(x_n) = 0$ oder $f(x_n) = g(x_n)$. OBdA fordern wir $f(1) = f(-1) = 0$, um die Funktion eindeutig zu beschreiben. Da wir die Funktion im Nullpunkt betrachten, wären auch andere Definitionen, statt der Festlegung der 2 Funktionswerte, möglich. Offensichtlich gilt $f(0) = 0$.



$$\begin{aligned}
N_{\text{epi } f}^{ko}(0, 0) &= N_{\text{epi } f}^F(0, 0) &= \{0\} \times \mathbb{R}_- \\
N_{\text{epi } f}^{\text{lim}, F}(0, 0) &= N_{\text{epi } f}^{Cl}(0, 0) &= \{(u, v) | v \leq -|u|\} \\
\partial^{ko} f(0) &= \partial^F f(0) &= \{0\} \\
\partial^{\text{lim}, F} f(0) &= \partial^{Cl} f(0) &= [-1, 1] \\
D^{Cl} f(0, 0)(v) &= \mathbb{R}, \quad D^F f(0, 0)(v) &= \{0 | -v \leq 0\} \\
D^{\text{lim}, F} f(0, 0)(v) &= \{u | -v \geq |u| \text{ oder } -v \leq -|u|\} \\
Cf(0, 0; u) &= \left[0, \frac{1}{2}|u|\right] \\
Tf(0, 0; u) &= [-|u|, |u|] \\
Df(0) &\text{ existiert nicht.}
\end{aligned}$$

Eine analoge Konstruktion kann man zu einer anderen Funktion machen, wodurch die Funktion im Punkte $(0,0)$ sogar differenzierbar wird.

Beispiel 6 (Sägezahnfunktion $f(x)$ zu $g(x) = x^2$) Sei $f : \mathbb{R} \rightarrow \mathbb{R}_+$ analog zu Beispiel 5 definiert. Offensichtlich gilt wieder $f(0) = 0$.

$$\begin{aligned}
N_{\text{epi } f}^{ko}(0, 0) &= N_{\text{epi } f}^F(0, 0) &= \{0\} \times \mathbb{R}_- \\
N_{\text{epi } f}^{\text{lim}, F}(0, 0) &= N_{\text{epi } f}^{Cl}(0, 0) &= \{(u, v) | v \leq -|u|\} \\
\partial^{ko} f(0) &= \partial^F f(0) &= \{0\} \\
\partial^{\text{lim}, F} f(0) &= \partial^{Cl} f(0) &= [-1, 1] \\
D^{Cl} f(0, 0)(v) &= \mathbb{R}, \quad D^F f(0, 0)(v) &= \{0\} \\
D^{\text{lim}, F} f(0, 0)(v) &= \{u | -v \geq |u| \text{ oder } -v \leq -|u|\} \\
Cf(0, 0; u) &= \{0\} \\
Tf(0, 0; u) &= [-|u|, |u|] \\
Df(0)u &= 0
\end{aligned}$$

Die Fréchet-Richtungsableitung fällt hier mit der Kontingent-Ableitung zusammen, was nach Definition zu erwarten ist. Die Ableitung existiert für die Funktion zwar im Nullpunkt aber $\partial^{\text{lim}, F} f(0)$ und $\partial^{Cl} f(0)$ enthalten zusätzliche Elemente. Interessant ist zudem, dass sich (bis auf den Wegfall der Vorzeichenbeschränkung bei $D^F f(0, 0)(v)$) nur die

Kontingient-Ableitung erheblich geändert hat im Vergleich zur anderen Sägezahnfunktion, die dieser ja recht ähnlich ist.

Nun stelle ich eine sehr einfache, stückweise lineare, Funktion vor, deren Definitionsbereich einmal nicht eindimensional ist.

Beispiel 7 (Himmel-und-Hölle-Funktion) Dies ist eine Funktion $F : \mathbb{R}^2 \rightarrow \mathbb{R}$, welche auf der y -Achse und der x -Achse konstant Null ist, vom Nullpunkt aus in die 4 Richtungen mit $|x| = |y|$ den Anstieg Eins hat und ansonsten durch lineare Fortsetzung beschrieben ist.

Zur konkreten Beschreibung verwenden wir die zwei linearen Hilfsfunktionen

$$L_1(x, y) = y \text{ und } L_2(x, y) = x.$$

Nun können wir umformulieren:

$$F(x, y) = \begin{cases} L_1(x, y), & 0 \leq y \leq |x| \\ L_2(x, y), & 0 \leq x \leq |y| \\ -L_2(x, y), & 0 \geq x \geq -|y| \\ -L_1(x, y), & 0 \geq y \geq -|x| \end{cases} = \begin{cases} |y|, & |y| \leq |x| \\ |x|, & |x| \leq |y| \end{cases} = \min \{|x|, |y|\}$$

und berechnen

$$\begin{aligned} N_{\text{epi } F}^{ko}(0, 0) &= N_{\text{epi } F}^F(0, 0) = \{0\}^2 \times \mathbb{R}_-, \\ N_{\text{epi } F}^{lim, F}(0, 0) &= \{(u, v, t) \mid uv = 0, t \leq -|u| - |v|\}, \\ N_{\text{epi } F}^{Cl}(0, 0) &= \{(u, v, t) \mid t \leq -|u| - |v|\} \\ \partial^{ko} F(0) &= \partial^F f(0) = \{0\}, \\ \partial^{lim, F} F(0) &= ([-1, 1] \times \{0\}) \cup (\{0\} \times [-1, 1]), \\ \partial^{Cl} F(0) &= \{(u, v) \mid |u| + |v| \leq 1\} \\ D^F F(0, 0)(v) &= \{(0, 0) \mid -v \leq 0\}, \\ D^{lim, F} F(0, 0)(v) &= \{u \mid u_1 u_2 = 0, v \geq |u_1| + |u_2|\} \cup \{u \mid -v = |u_1| + |u_2|\}, \\ D^{Cl} F(0, 0)(v) &= \mathbb{R}^2, \\ CF(0, 0; u) &= \{F'(x; u)\} = \{\min \{|u_1|, |u_2|\}\} = \{F(u)\}, \\ TF(0, 0; u) &= [-\max \{|u_1|, |u_2|\}, \max \{|u_1|, |u_2|\}] \\ DF(0) &\text{ existiert nicht} \\ N_{\text{gph } F}^{lim, F}(0, 0, 0) &= \{(u_1, u_2, v) \mid u_1 u_2 = 0, -v \geq |u_1| + |u_2|\} \\ &\cup \{(u_1, u_2, v) \mid v = |u_1| + |u_2|\}. \end{aligned}$$

Der Name kommt vom Aussehen des Graphs, welcher an ein Kinderspiel erinnert. Da das zugehörige Faltobjekt auch Schnapper genannt wird, verwende ich diesen Begriff.

Hier unterscheiden sich alle Ableitungen sehr stark. Das Fréchet-Subdifferential ist einwertig, aber die Funktion ist nicht differenzierbar. Auch ansonsten ist diese Funktion günstig um bestimmte Gesetzmäßigkeiten zu prüfen, wie wir später sehen werden.

Alle diese Beispiele lassen erwarten, dass eine Berechnung der gewünschten Größen sehr schwierig sein kann. Grenzwertverhalten lässt sich schwer voraussagen und selbst bei einfachen Funktionen wie $-|x|$ unterscheiden sich die verschiedenen Ableitungsbegriffe sehr stark.

Während die bisher genannten Funktionen schon interessante Werte lieferten, so deutet die folgende an, wie aufwendig die Betrachtung selbst von Lipschitzfunktionen sein kann.

Beispiel 8 (Blitzfunktion) Die Konstruktion dieser Funktion $f : [0, 1] \rightarrow \mathbb{R}$ wird im [KK02] als erstes „Basic example“ beschrieben. Ich möchte mich damit begnügen ein paar Eigenschaften aufzuzählen.

f hat Lipschitzrang 1, ist auf einer dichten Menge nicht differenzierbar, hat auf dichten Mengen Maxima bzw. Minima und das konstante Clarke-Subdifferential $\partial^{Cl} f(x) = [-1, 1]$.

Leichte Variationen der Konstruktion liefern Funktionen mit den selben Eigenschaften aber ganz anderem Aussehen. Dies zeigt, dass das Clarke-Subdifferential so gut wie keine brauchbare Information über eine Funktion enthalten kann. Außerdem gilt zwar, dass die Menge der Punkte, in denen eine Lipschitzfunktion nicht differenzierbar ist, Maß Null haben, aber trotzdem dicht sein kann. Auch das Finden von Differenzierbarkeitspunkten kann sehr schwer werden.

Es folgen ein paar einfache Beispiele, die die „Zusammenhänge“ veranschaulichen zwischen Teilmengenbeziehungen und zugehörigen Normalenkegeln.

Beispiel 9 (positiver Orthant)

$$\begin{aligned} M_1 &:= \{(x, y) | x, y \geq 0\} = \mathbb{R}_+^2 \\ N_{M_1}^\alpha(0, 0) &= \{(u, v) | u, v \leq 0\} = \mathbb{R}_-^2 \quad \forall \alpha \in \{ko; F; lim, F; Cl\} \end{aligned}$$

Beispiel 10 (Rand des positiven Orthanten)

$$\begin{aligned} M_2 &:= \{(x, y) | xy = 0, (x, y) \geq 0\} = \text{Rand}(M_1) \\ N_{M_2}^{ko}(0, 0) = N_{M_2}^F(0, 0) &= \mathbb{R}_-^2, \\ N_{M_2}^{lim, F}(0, 0) &= \mathbb{R}_-^2 \cup M_2, \\ N_{M_2}^{Cl}(0, 0) &= \mathbb{R}^2 \end{aligned}$$

Beispiel 11 (x -Achse)

$$\begin{aligned} M_3 &:= \{(x, y) | x = 0, y \geq 0\} \subset M_2 \\ N_{M_3}^\alpha(0, 0) &= \mathbb{R} \times \mathbb{R}_- \quad \forall \alpha \in \{ko; F; lim, F; Cl\} \end{aligned}$$

Beispiel 12 (Koordinatenachsen)

$$\begin{aligned} M_4 &:= \{(x, y) | xy = 0\} = M_2 \cup -M_2 \supset M_2 \\ N_{M_4}^{ko}(0, 0) = N_{M_4}^F(0, 0) &= \{0\}, \\ N_{M_4}^{lim, F}(0, 0) &= M_4, \\ N_{M_4}^{Cl}(0, 0) &= \mathbb{R}^2 \end{aligned}$$

4 Allgemeine Sätze

Damit diese Arbeit übersichtlich bleibt, fasse ich Aussagen, die ich für wichtig halte oder die ich für spätere Beweise benötige, in diesem Kapitel zusammen. Es werden ein paar Eigenschaften der zuvor genannten Definitionen gezeigt, aber es wird noch nicht auf die Zusammenhänge oder tiefere Aussagen eingegangen. Wer etwas mehr über die vorangegangenen Definitionen wissen will, wird hier vielleicht fündig. Jedem anderen Leser ist es zu empfehlen, diesem Bereich zum Nachschlagen bei Verweisen aus späteren Kapiteln zu benutzen.

4.1 Ekeland's Variationsprinzip

Ein sehr wichtiger Satz der Optimierung ist

Satz 4.1 (Ekeland's Variationsprinzip) Sei $f : M \rightarrow \overline{\mathbb{R}}$ unterhalbstetig mit abgeschlossener Menge $M \subset \mathbb{R}^n$ und $v := \inf_{x \in M} f(x) \in \mathbb{R}$ und x_ϵ sei ϵ -optimal, d.h. $f(x_\epsilon) \leq v + \epsilon$. Dann gilt $\forall \alpha > 0 \exists z_\alpha \in B(x_\epsilon, \alpha) \cap M$:

$$f(z_\alpha) \leq f(x) + \frac{\epsilon}{\alpha} d(x, z_\alpha) \quad \forall x \in M \quad \text{und} \quad f(z_\alpha) \leq f(x_\epsilon).$$

Beweis nach [KK02]: Man betrachte die Abbildung $h(z) := \inf_{x \in M} (f(x) + \frac{\epsilon}{\alpha} d(x, z))$ und folgere direkt mit der Dreiecksungleichung, dass h Lipschitzstetig ist mit Konstante $L := \frac{\epsilon}{\alpha}$, denn die Endlichkeit und Existenz folgt aus der Existenz von v :

$$f(x) + \frac{\epsilon}{\alpha} d(x, z_1) \leq f(x) + \frac{\epsilon}{\alpha} d(x, z_2) + \frac{\epsilon}{\alpha} d(z_1, z_2) \Rightarrow h(z_1) \leq h(z_2) + \frac{\epsilon}{\alpha} d(z_1, z_2).$$

Wenn $h(z_k) < f(z_k)$, finden wir $z_{k+1} \in M$ mit $f(z_{k+1}) + Ld(z_{k+1}, z_k) < f(z_k)$ und $f(z_{k+1}) + Ld(z_{k+1}, z_k) < h(z_k) + 2^{-k}$. Dies liefert

$$\begin{aligned} d(z_0, z_{k+1}) &\leq \sum_{s=0}^k d(z_s, z_{s+1}) < \frac{1}{L} \sum_{s=0}^k (f(z_s) - f(z_{s+1})) = \frac{\alpha}{\epsilon} (f(z_0) - f(z_{k+1})) \leq \alpha \\ &\Leftrightarrow f(z_0) - f(z_{k+1}) \leq \epsilon, \end{aligned}$$

weshalb wir mit $z_0 := x_\epsilon$ initialisieren.

Wenn $h(z_k) \geq f(z_k)$, wären wir somit fertig mit $z_k =: z_\alpha$. Da $f(z_k)$ fallend ist, können wir statt z_0 in der Ungleichung andere z_s einsetzen, was zeigt, dass wir eine Cauchy-Folge mit Grenzwert z_α haben. $z_\alpha \in B(x_\epsilon, \alpha) \cap M$ und $f(z_\alpha) \leq f(x_\epsilon)$ sind bereits gezeigt, da sie für die z_k gelten und f unterhalbstetig ist. Nun gilt außerdem die letzte gewünschte Bedingung:

$$f(z_\alpha) \leq \liminf_{k \rightarrow \infty} f(z_{k+1}) \leq \liminf_{k \rightarrow \infty} (f(z_{k+1}) + Ld(z_{k+1}, z_k)) \leq \liminf_{k \rightarrow \infty} (h(z_k) + 2^{-k}) = h(z_\alpha).$$

□

Führen wir die Größe $\delta = \frac{\epsilon}{\alpha}$ ein, so wird obiges z_α auch δ -Ekeland-Punkt genannt und lokaler δ -Ekeland-Punkt, wenn die Bedingung nur für alle zulässigen x nahe z_α gilt.

4.2 Konvexe Analysis

Zur Vollständigkeit wiederhole ich die grundlegenden Definitionen.

Definition 4.2 (konvex) Eine Menge M heißt konvex genau dann, wenn

$$\forall x, y \in M : tx + (1 - t)y \in M \quad \forall t \in (0, 1).$$

Eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ heißt genau dann konvex, wenn $\text{epi } f \subset \mathbb{R}^n \times \mathbb{R}$ konvex ist.

Die konvexe Analysis ist meist die Grundlage der verallgemeinerten Ableitungen, insbesondere der Subdifferentiale. Zudem gibt es in der konvexen Analysis sehr viele schöne Sätze, die man für die Verallgemeinerungen nutzt, oder deren Aussage man auch für den nichtkonvexen Fall haben will. Die **Trennungssätze** der konvexen Analysis sind sehr bekannt. Ich werde sie nicht wiederholen, da man sie in jedem guten Funktionalanalysisbuch findet und sie meist in der Analysis-Grundvorlesung hört. Es sollte hier nur kurz erinnert werden, dass für konvexe Funktionen das lokale Minimum gleichzeitig das globale ist.

Nun werde ich einige elementare Sätze aufzählen, ohne sie zu beweisen. Im folgenden gilt $f : X \rightarrow \bar{\mathbb{R}}$ und $\text{dom } f$ bezeichnet die Menge $\{x \in X | f(x) < \infty\}$.

Satz 4.3 Sei $X = \mathbb{R}^n$, $x \in \text{int}(\text{dom}(f))$ und f konvex, dann ist f um x lokal lipschitzstetig.

Dies zeigt, dass lokale Lipschitzstetigkeit eine gar nicht so außergewöhnliche Forderung an eine Funktion ist, deren Minimum wir bestimmen wollen.

Satz 4.4 Sei $X = \mathbb{R}^n$, $x \in \text{int}(\text{dom}(f))$ und f konvex, dann gilt $\partial^{ko} f(x) \neq \emptyset$.

Der Beweis erfolgt über einen Trennungssatz und Fallunterscheidung.

Dieser Satz ist unter anderem eine Motivation für das Clarke- und das Limiting-Fréchet-Konzept, denn durch die dort zugrundeliegende Grenzwertbildung erzeugt man, auch wenn es nicht sofort ersichtlich ist, nichtleere Subdifferentiale bei nicht-konvexen lipschitzstetigen Funktionen.

Definition 4.5 (konvexe Hülle) Mit $\text{conv}(M)$ bezeichnen wir die konvexe Hülle einer Menge, das heißt, die kleinste konvexe Menge die M enthält bzw. die Menge aller Konvexkombinationen aus Elementen von M .

Satz 4.6 (Caratheodory) Sei $M \subset \mathbb{R}^n$. Dann gilt,

$$\text{conv}(M) = \left\{ \sum_{i=0}^{\mathbf{n}} \lambda_i m_i \mid m_i \in M, \lambda_i \geq 0 \quad \forall i, \sum_{i=0}^{\mathbf{n}} \lambda_i = 1 \right\}.$$

Offensichtlich ändert die konvexe Hülle nichts bei konvexen Mengen. Ein großer Vorteil konvexer Kegel ist, dass man recht einfach mit ihnen rechnen kann.

Folgerung 4.7 (konvexe Kegel) Gelte $\emptyset \neq A \subset \mathbb{R}^n$. Dann definiert $\mathbb{R}_+ A$ einen beliebigen Kegel. Seien B und C zwei weitere Kegel. Dann gilt:

$$\text{conv}(\mathbb{R}_+ A) = \mathbb{R}_+ A + \mathbb{R}_+ A = \mathbb{R}_+ \text{conv}(A) \quad \text{und} \quad \text{conv}(B + C) = \text{conv}(B) + \text{conv}(C).$$

4.3 Fréchet-Konzept

Das Subdifferential $\partial^F f(x)$ kann man über den Normalenkegel definieren, allerdings ergibt sich für lokal lipschitzstetige Funktionen eine Formel, mit der man wesentlich besser arbeiten kann und die sehr oft stillschweigend benutzt wird.

Satz 4.8 (Fréchet-Subdifferential) Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ um x lokal lipschitzstetig. Dann gilt

$$\partial^F f(x) = \{v \mid \langle v, y - x \rangle \leq f(y) - f(x) + o(y - x) \ \forall y \in \mathbb{R}^n\}. \quad (4.1)$$

Beweis: Es sind allerdings nur y nahe x interessant. Zunächst nehmen wir die allgemeine Definition und verschärfen sie auf einen „Sonderfall“. Danach ersetzen wir den o -Term.

$$\begin{aligned} \partial^F f(x) &= \{v \mid \langle v, y - x \rangle \leq t - f(x) + o(y - x), t \geq f(y)\} \\ &= \{v \mid \langle v, y - x \rangle \leq f(y) - f(x) + o(y - x), f(y) - f(x)\} \\ &= \{v \mid \langle v, y - x \rangle \leq f(y) - f(x) + o(y - x)\}. \end{aligned}$$

Im ersten Schritt haben wir benutzt, dass die Bedingung auch für $t = f(y)$ gelten muss und die Abschätzung dann am schärfsten ist, da $t \geq f(y)$ und der o -Term nur klein wird, wenn sich t dem $f(x)$ nähert, was parallel geschieht mit der Näherung von y an x und somit nähert sich t auch $f(y)$.

Im zweiten Schritt benutzen wir die lokale Lipschitzstetigkeit. Sei L die zugehörige Konstante und y oBdA in einer hinreichend kleinen Umgebung für diese lokale Konstante, aber sonst beliebig. Nun gilt

$$\frac{o(y - x, f(y) - f(x))}{\|(y - x, f(y) - f(x))\|} = \frac{o(y - x, f(y) - f(x))}{\|y - x\| + \|f(y) - f(x)\|} \geq \frac{o(y - x, f(y) - f(x))}{(1 + L) \|y - x\|} \geq 0.$$

Grenzwertbildung liefert die notwendige Beziehung der o -Terme. Welche Norm hier verwendet wird ist egal, da die Normen im \mathbb{R}^n alle äquivalent sind. \square

Im folgenden leite ich ausgehend von elementaren Eigenschaften mit einer Folge von Sätzen eine Teilmengenbeziehung für $N_M^F(x)$ bei einer recht einfachen Beschreibung von M her, die später verwendet wird zum Vergleich mit dem Limiting-Fréchet-Konzept.

Satz 4.9 (abgeschlossene konvexe Kegel) Sowohl der Fréchet-Normalenkegel, der Kontingent-Kegel, der Normalenkegel der konvexen Optimierung als auch der Clarke-Normalenkegel sind

- abgeschlossene,
- konvexe
- Kegel.

Beweis: Die Beweise folgen direkt aus den Definitionen. \square

Satz 4.10 (Teilmengenbeziehung) Der Fréchet Normalenkegel erfüllt die Bedingung

$$A \subset B \subset \mathbb{R}^n \Rightarrow N_A^F(x) \supset N_B^F(x) \ \forall x \in A.$$

Beweis: Folgt direkt aus der Definition. \square

Bemerkung Die Forderung $x \in A$ ist hier wichtig und verhindert einen Übergang zum „lim sup“ für den Limiting-Fréchet-Normalenkegel! Für den Normalenkegel der konvexen Optimierung und den Kontingent-Kegel ließe sich auch eine einfache Teilmengeneigenschaft zeigen, welche wir für unsere Betrachtungen allerdings nicht brauchen. Sowohl für den Clarke- als auch den Limiting-Fréchet-Normalenkegel gibt es keine solche Beziehung, wie man leicht an den Beispielen am Ende des Beispielabschnitts nachprüft.

Folgerung 4.11 Sei $M = P \cap Q$, dann gilt

$$N_M^F(x) \supset N_P^F(x) \text{ und } N_M^F(x) \supset N_Q^F(x) \quad \forall x \in M.$$

Da $N_M^F(x)$ jedoch ein konvexer Kegel ist, folgt sofort mit 4.7

$$N_M^F(x) \supset N_P^F(x) + N_Q^F(x) \quad \forall x \in M.$$

Satz 4.12 Sei $g : \mathbb{R}^n \rightarrow \mathbb{R}^q$ stetig differenzierbar, $Q \subset \mathbb{R}^q$, $M = g^{-1}(Q)$, dann gilt

$$N_M^F(x) \supset Dg^T(x)N_Q^F(g(x)) \quad \forall x \in M.$$

Beweis: Sei $u \in N_Q^F(g(x))$ und $y \in g^{-1}(Q) = M$ beliebig. Dann gilt

$$\begin{aligned} \langle Dg(x)^T u, y - x \rangle &= \langle u, Dg(x)(y - x) \rangle \\ &= \langle u, g(y) - g(x) + o(y - x) \rangle, \quad \text{nach Definition von } Dg \\ &\leq \langle u, g(y) - g(x) \rangle + o(y - x), \quad \text{weil } u \text{ stetig linear ist} \\ &\leq o(g(y) - g(x)) + o(y - x), \quad \text{denn } u \in N_Q(g(x)) \\ &\leq o(Dg(x)(y - x)) + o(y - x), \quad \text{da } o(o(y - x)) \leq o(y - x) \\ &\leq o(y - x), \quad \text{weil } Dg(x) \text{ stetig linear ist.} \end{aligned}$$

Aufgrund der beliebigen Wahl von $y \in M$ folgt $Dg(x)^T u \in N_M^F(x)$. \square

Bemerkung Damit M abgeschlossen ist, sollte man Q abgeschlossen wählen und im folgenden Satz auch N .

Folgerung 4.13 Sei $N \subset \mathbb{R}^n$ und $M = N \cap g^{-1}(Q)$, dann gilt

$$N_M^F(x) \supset Dg(x)^T N_Q^F(g(x)) + N_N^F(x) \quad \forall x \in M.$$

Den Weg von Fréchet- zu Limiting-Fréchet-Normalenkegel über Grenzwertbildung kennen wir schon aus der Definition, aber er überträgt sich auch auf die Subdifferenziale.

Satz 4.14 ($\partial^F f(x)$ vs. $\partial^{lim,F} f(x)$)

$$\limsup_{x_k \xrightarrow{\text{in } M} x} \partial^F f(x_k) = \partial^{lim,F} f(x)$$

Beweis: Das Subdifferential ergibt sich aus dem Normalenkegel an $\text{epi } f$, indem wir alle Elemente betrachten mit letzter Komponente -1 . Bilden wir nun Grenzwerte für den Übergang von Fréchet zu Limiting-Fréchet, so muss die letzte Komponente natürlich oBdA für alle Elemente der Folge schon negativ sein. Nun können wir während der Grenzwertbildung so skalieren, dass die letzte Komponente -1 bleibt. Der Skalierungsfaktor konvergiert dann gegen Eins. Folglich genügt es, Elemente mit letzter Komponente -1 für die Grenzwertbildung zu betrachten. \square

4.4 Clarke-Konzept

Wie bereits bei der Definition des zugehörigen Tangentialkegels erwähnt, gibt es auch andere Zugänge zum Clarke-Subdifferential. Dieser wird hier angedeutet und der Bezug zu Clarkes verallgemeinerten Ableitung.

Satz 4.15 (Clarke-Richtungsableitung) $f^{Cl}(x; \cdot)$ ist:

- endlich, wenn L eine lokale Lipschitzkonstante von f um x ist, mit

$$|f^{Cl}(x; u)| \leq L \|u\|,$$

- positiv homogen, d.h.

$$f^{Cl}(x; \lambda^+ u) = \lambda^+ f^{Cl}(x; u) \quad \forall \lambda^+ \in \mathbb{R}_+,$$

- subadditiv, d.h.

$$f^{Cl}(x; u + v) \leq f^{Cl}(x; u) + f^{Cl}(x; v), \quad (4.2)$$

- und folglich auch konvex (in der zweiten Komponente).

Beweis: Alles folgt sofort aus der Definition. \square

Aufgrund der eben gezeigten Konvexität bietet es sich an, das klassische Subdifferential der konvexen Optimierung anzuwenden.

Satz 4.16 (Clarke-Subdifferential) Es gilt:

$$\partial^{Cl}(f)(x) = \partial^{ko}(f^{Cl}(x; \cdot))(0) \quad (4.3)$$

Sei f lokal lipschitzstetig um $x \in \mathbb{R}^n$, dann gilt

$$\partial^{Cl}(f)(x) = \text{conv} \{v = \lim Df(x_k) | x_k \rightarrow x, Df(x_k) \text{ existiert}\} \quad (4.4)$$

Dies sind beides nichttriviale Aussagen. Die zugehörigen Beweise sind in [Cla83] Korollar S.61 und Theorem 2.5.1 zu finden. Da wir das Subdifferential bzgl. der Ableitungsrichtungen bilden, erhalten wir eher eine Art lineare Näherung der Clarke-Richtungsableitung. Dies hat nichts mit einer zweiten Ableitung zu tun.

Folgerung 4.17 Aus (4.3), Satz 4.15 und Satz 4.4 folgt $\partial^{Cl}f(x) \neq \emptyset$, wenn f lokal um x lipschitzstetig ist. Alternativ kann man zum Beweis Gleichung (4.4) und den Satz von Rademacher benutzen.

Der Vollständigkeit halber, erwähne ich hier, dass nach Definition gilt

$$f^{Cl}(x; u) = \sup \{v \in Tf(x; u)\}.$$

5 Teilmengenbeziehungen

5.1 Vergleich der Normalenkegelkonzepte

Nun kommen wir zu einem ersten konkreten Vergleich der verschiedenen Definitionen. Hierbei liegt der Fokus bei den Normalenkegelkonzepten, da sich die Beziehungen für die Subdifferenziale und Co-Differenziale daraus sofort übertragen.

Satz 5.1 (einfache Beziehungen) Direkt aus den Definitionen folgt

$$\begin{aligned} N_M^{ko}(x) &\subset N_M^F(x), \\ N_M^B(x) &\subset N_M^{Cl}(x), \quad \text{da } T_M^B(x) \supset T_M^{Cl}(x) \text{ und} \\ N_M^F(x) &\subset N_M^{lim,F}(x). \end{aligned}$$

Satz 5.2 (Polarkegel) Mit $A \subset X = \mathbb{R}^n$ gilt $A = (A^*)^*$ genau dann, wenn A ein abgeschlossener, konvexer Kegel ist.

Beweisskizze: Die Richtung „ \Rightarrow “ folgt direkt aus der Beobachtung, dass der Polarkegel immer ein abgeschlossener, konvexer Kegel ist. Die andere Richtung „ \Leftarrow “ wird mit dem Trennungssatz für konvexe Mengen gezeigt, nachdem die Beziehung $A \subset (A^*)^*$ einfach nachgeprüft wurde. Es würde reichen zu fordern, dass X ein reflexiver Banachraum ist, da dort auch der notwendige Trennungssatz gilt. \square

Diese Aussage ist insofern interessant, als dass sie eine Eins-zu-Eins-Beziehung zwischen Tangential- und Normalenkegel unter bestimmten Voraussetzungen liefert. Diese erfüllt allerdings nur der Clarke-Tangentialkegel immer.

Satz 5.3 (Bouligand- und Fréchet-Normalenkegel) Es gilt generell $N_M^F(x) \subset N_M^B(x)$ und Gleichheit, da $X = \mathbb{R}^n$.

Beweisskizze: Seien $t_k \searrow 0, u_k \rightarrow u \in T_M^B(x)$ beliebig mit $x + t_k u_k \in M$.

$$\langle v, x - (x + t_k u_k) \rangle = \langle v, t_k u_k \rangle \leq o(t_k u_k) \Rightarrow \left\langle v, \frac{u_k}{\|u_k\|} \right\rangle \leq \frac{o(t_k u_k)}{\|t_k u_k\|} \rightarrow 0$$

$$\Rightarrow N_M^F(x) \subset T_M^B(x)^*.$$

Angenommen die andere Teilmengenbeziehung gilt nicht für ein $v \in T_M^B(x)^*$ und sei $x_k \xrightarrow{M} x$ eine zugehörige Folge mit $\langle v, x_k - x \rangle > o(x_k - x)$, von der wir nun eine passend Teilfolge betrachten.

$$t_k := \|x_k - x\| \searrow 0, u_k := \frac{x_k - x}{\|x_k - x\|} \xrightarrow{oBdA} u \in T_M^B(x), \text{ da } X = \mathbb{R}^n \text{ und } x_k = x + t_k u_k$$

$$v \in T_M^B(x)^* \Rightarrow 0 \geq \langle v, u \rangle = \lim_{k \rightarrow \infty} \frac{\langle v, x_k - x \rangle}{\|x_k - x\|} \Rightarrow \langle v, x_k - x \rangle \leq o(x_k - x)$$

$\Rightarrow N_M^F(x) \supset T_M^B(x)^*$ durch Widerspruch, da auch eine Teilfolge von $x_k \xrightarrow{M} x$ zeigen müsste, dass $v \notin N_M^F(x)$ gilt über die Definition. \square

Satz 5.4 (Limiting-Fréchet- und Clarke-Normalenkegel) Aus [Cla83] Satz 2.5.7 folgt:

$$N_M^{Cl}(x) = \text{conv}(N_M^{lim,F}(x)) \quad (5.1)$$

und somit $N_M^{lim,F}(x) \subset N_M^{Cl}(x)$,

da $X = \mathbb{R}^n$. Der zugehörige Beweis ist dort sehr aufwändig und verwendet sehr viele Sätze zum Clarke-Konzept und andere Darstellungen der beiden Normalenkegel.

Folgerung 5.5 ($\partial^{Cl} f(x)$ vs. $\text{conv}\{\partial^{lim,F} f(x)\}$) Es gilt generell die Beziehung

$$\partial^{Cl} f(x) \supset \text{conv}\{\partial^{lim,F} f(x)\}$$

Falls $N_{\text{epi } f}^{Cl}(x) \cap \mathbb{R}^n \times \{0\} = \emptyset$ erfüllt ist, gilt sogar

$$\partial^{Cl} f(x) = \text{conv}\{\partial^{lim,F} f(x)\}.$$

Die zusätzliche Forderung sichert, dass der Schnitt mit Elementen mit -1 in letzter Komponente und Bildung der konvexen Hülle vertauscht werden können. Bei Lipschitzfunktionen ist diese Bedingung erfüllt.

In meinem eigenen einfachen Beweis verwende ich nur lokale Lipschitzstetigkeit und weder die Hilfsbedingung noch den vorangegangenen aufwendigen Satz. Dafür verwende ich die Formel für das Fréchet-Subdifferential einer differenzierbaren Funktion, welche ich erst später beweise.

Eigener Beweis: Zunächst gilt

$$\begin{aligned} \partial^{Cl}(f)(x) &\stackrel{(4.4)}{=} \text{conv}\{v = \lim Df(x_k) \mid x_k \rightarrow x, Df(x_k) \text{ existiert}\} \\ &\stackrel{\text{Satz 10.1}}{=} \text{conv}\{v = \lim v_k \mid v_k \in \partial^F f(x_k), x_k \rightarrow x, Df(x_k) \text{ existiert}\} \\ &\subset \text{conv}\{\partial_M^{lim,F} f(x)\}. \end{aligned}$$

Andererseits zeigt eine Diagonalisierungsargumentation angewendet auf (4.4), da f lokal lipschitzstetig ist, dass $\partial^{Cl}(f)(x)$ eine abgeschlossene Abbildung ist. Damit zeigen wir nun die Umkehrung

$$\begin{aligned} \text{conv}\{\partial^{lim,F} f(x)\} &= \text{conv}\left\{\limsup_{x_k \rightarrow x} \partial^F f(x_k)\right\} && \text{(Ende von Abschnitt 4.3)} \\ &\subset \text{conv}\left\{\limsup_{x_k \rightarrow x} \partial^{Cl} f(x_k)\right\} && (N_M^F(y) = N_M^B(y) \subset N_M^{Cl}(y)) \\ &= \text{conv}\{\partial^{Cl} f(x)\} && \text{(Abgeschlossenheit)} \\ &= \partial^{Cl} f(x) && (N_M^{Cl}(y) \text{ ist konvex}). \end{aligned}$$

□

Folgerung 5.6 (Co-Differential) Es folgt analog

$$D^{Cl}f(x, f(x))(v) \supset \text{conv } D^{lim,F}f(x, f(x))(v).$$

Wie man an den Beispielen in Abschnitt 3 sieht, ist Gleichheit noch nicht einmal bei der einfachen Betragsfunktion gegeben. Bei Abbildungen $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ oder mehrwertigen Abbildungen bleibt die Teilmengenbeziehung natürlich erhalten.

Satz 5.7 (Zusammenfassung) Mit $M \subset X = \mathbb{R}^n$ haben wir nun insgesamt:

$$N_M^{ko}(x) \subset N_M^F(x) = N_M^B(x) \subset N_M^{lim,F}(x) \subset \text{conv}(N_M^{lim,F}(x)) = N_M^{Cl}(x).$$

Analoges ergibt sich auch für die zugehörigen Subdifferenziale und die Co-Differenziale.

Satz 5.8 (konvexer Fall) Wenn M konvex ist, so sind alle Normalenkegel gleich.

Analoges gilt für die Subdifferenziale konvexer Funktionen.

Beweis: Es genügt aufgrund der Teilmengenbeziehungen und der Polarbedingung zu zeigen, dass $y - x \in T_M^{Cl}(x)$ gilt für $y \in M$.

Betrachte dazu beliebige $x_k \xrightarrow{M} x$ und $t_k \searrow 0$ wobei oBdA $t_k < 1$ gelte.

Dann ist $x_k + (1 - t_k)(y - x_k) \in M$, da M konvex. Dies liefert die nötige Bedingung in der Definition von $T_M^{Cl}(x)$ mit $v_k = y - x_k \rightarrow v = y - x$. \square

5.2 Beziehungen zwischen Normalenkegel und Subdifferential

Bisher haben wir schon die Beziehung und Definition

$$v \in \partial^\alpha f(x) \Leftrightarrow (v, -1) \in N_{epi F}^\alpha.$$

Für die „duale“ Version brauchen wir die Indikatorfunktion der Menge M .

Definition 5.9 (Indikatorfunktion)

$$i_M(x) = \begin{cases} 0 & , x \in M \\ \infty & , x \notin M \end{cases}$$

Nun haben wir sofort die andere bekannte Beziehung

$$\partial^\alpha i_M(x) = N_M^\alpha(x),$$

wobei hier die später gezeigte Kreuzproduktregel 6.1 angewandt wird auf $epi i_M = M \times \mathbb{R}_+$.

Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ lokal lipschitzstetig und

$$M = \{x | f(x) \leq 0\}.$$

Mit Gleichung (4.1) und wegen $N_M^F(x) = \mathbb{R}_+ N_M^F(x)$ erhält man den Satz:

Satz 5.10 ($\partial^F f(x)$ vs. $N_M^F(x)$)

$$\mathbb{R}_+ \partial^F f(x) \subset N_M^F(x) \quad \forall x : f(x) = 0.$$

In den anderen Fällen ist der Normalenkegel leer oder besteht nur aus 0, was für das Subdifferential im Allgemeinen nicht gilt.

Da die Aussage nur für spezielle x gilt, ist ein Grenzwertübergang nicht möglich. Außerdem zeigt die Betragsfunktion, dass dies eine echte Inklusion ist. Bei zusätzlichen Voraussetzungen erhält man die umgekehrte Inklusion für das Clarke-Subdifferential vgl. [Cla83] Satz 2.9.10. In der folgenden Version hab ich die Voraussetzungen vereinfacht, da ich zum Beispiel von einer lokal lipschitzstetigen Funktion ausgehe.

Satz 5.11 ($\partial^{Cl} f(x)$ vs. $N_M^{Cl}(x)$) Haben wir $f(x) = 0$ und zusätzlich $0 \notin \partial^{Cl} f(x)$, so gilt

$$\mathbb{R}_+ \partial^{Cl} f(x) \supset N_M^{Cl}(x).$$

Beweisanmerkung: Den Beweis kann man in [Cla83] nachschlagen oder durch den nachfolgenden Satz beweisen, indem man zur konvexen Hülle übergeht. \square

In [RW98] S. 424 Satz 10.3 findet man die entsprechende Aussage für $N_M^{lim,F}(x)$.

Satz 5.12 ($\partial^{lim,F} f(x)$ vs. $N_M^{lim,F}(x)$) Wir fordern wieder $f(x) = 0$ und $0 \notin \partial^{lim,F} f(x)$ und erhalten

$$\mathbb{R}_+ \partial^{lim,F} f(x) \supset N_M^{lim,F}(x).$$

Beweisskizze: Man sieht sofort

$$M \times \{0\} = (\mathbb{R}^n \times \{0\}) \cap \text{epi } f.$$

Um Satz 8.4 anzuwenden, welcher später folgt, muss man also $N_{\text{epi } f}^{lim,F}(x, 0)$ untersuchen. Aus der Definition von $N_{\text{epi } f}^{lim,F}$ über $N_{\text{epi } f}^F$ erkennt man, dass $N_{\text{epi } f}^{lim,F}(x, 0) \cap (\mathbb{R}^n \times (0, \infty)) = \emptyset$ und wenn L eine lokale Lipschitzkonstante von f um x ist, zeigt man leicht

$$\forall (v, -t) \in N_{\text{epi } f}^F(y), t > 0, y \text{ nahe } (x, 0) \text{ gilt } \|v\| \leq Lt.$$

Geht man zum Grenzwert $t \searrow 0$ über, so folgt, dass $N_{\text{epi } f}^{lim,F}(x, 0) \cap (\mathbb{R}^n \times \{0\}) = \{0\}$ ist. Somit ergibt sich

$$N_{\text{epi } f}^{lim,F}(x, 0) = \mathbb{R}_+ (N_{\text{epi } f}^{lim,F}(x, 0) \cap \mathbb{R}^n \times \{-1\}) = \mathbb{R}_+ (\partial^{lim,F} f(x) \times \{-1\}).$$

Die Regularitätsbedingung in Satz 8.4 lautet hier $0 \notin \partial^{lim,F} f(x)$ und die erhaltene Teilmengenbeziehung ist

$$N_{M \times \{0\}}^{lim,F} \subset \{0\}^n \times \mathbb{R} + \mathbb{R}_+ (\partial^{lim,F} f(x) \times \{-1\}) = (\mathbb{R}_+ \partial^{lim,F} f(x)) \times \mathbb{R}.$$

Der Rest folgt über die Rechenregeln zum Normalenkegel vom Kreuzprodukt (vgl. Satz 6.1 weiter hinten). \square

Man bemerkt dabei, dass die Bedingung $0 \notin \partial^{Cl} f(x)$ im vorangegangenen Satz schärfer gewählt wurde als nötig. Allerdings ist das Clarke-Subdifferential oftmals und vor allem in unseren Anwendungen einfacher zu berechnen und somit die Bedingung in der Praxis wesentlich einfacher nachprüfbar.

Bemerkung (Regularität) Unter Regularität von f in x haben wir nun eine weitere Eins-zu-Eins-Beziehung zwischen Subdifferentialberechnung und Normalenkegelberechnung. Außerdem liefert uns diese Bedingung Regularität der Normalenkegel.

5.3 Beziehungen bzgl. Co-Differential, Kontingent- und Thibault-Ableitung

Satz 5.13 ($TF(x, y; u)$ vs. $CF(x, y; u)$) Sei $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ eine mehrwertige Abbildung. Die Beziehung der beiden Ableitungen lässt sich verschärfen zu

$$TF(x, y; u) \supset \limsup_{(x_i, y_i) \xrightarrow{\text{Gph } F} (x, y)} CF(x_i, y_i; u)$$

und es gilt außerdem, dass $TF(x, y; u)$ abgeschlossen ist als Abbildung

$$TF : \text{Gph } F \times \mathbb{R}^n \rightrightarrows \mathbb{R}^m.$$

Beweis: Da ohnehin aus der Definition $TF(x, y; u) \supset CF(x_i, y_i; u)$ folgt, genügt es die Abgeschlossenheit von TF zu zeigen. (Die Teilmengenbeziehung folgt dann mit festen u und $(x_i^k, y_i^k) = (x_i, y_i)$.)

Wir betrachten also

$$(x_i, y_i) \xrightarrow{\text{Gph } F} (x, y), (u_i, v_i) \rightarrow (u, v) \text{ mit } v_i \in TF(x_i, y_i; u_i)$$

Zu gegebenem n existiert deshalb ein i mit

$$\|(x_i, y_i) - (x, y)\| < \frac{1}{2n} \text{ und } \|(u_i, v_i) - (u, v)\| < \frac{1}{2n} \text{ weil } (x_i, y_i, u_i, v_i) \rightarrow (x, y, u, v).$$

Aufgrund der Definition der Thibault-Ableitung gilt:

$$\exists t_i^k \searrow 0, (x_i^k, y_i^k) \xrightarrow{\text{Gph } F} (x_i, y_i), (u_i^k, v_i^k) \rightarrow (u_i, v_i) \text{ mit } y_i^k + t_i^k v_i^k \in F(x_i^k + t_i^k u_i^k),$$

also insbesondere $(x_i^k, y_i^k, u_i^k, v_i^k, t_i^k) \rightarrow (x_i, y_i, u_i, v_i, 0)$, womit wir ein k wählen können mit

$$\|(x_i^k, y_i^k) - (x_i, y_i)\| < \frac{1}{2n}, \|(u_i^k, v_i^k) - (u_i, v_i)\| < \frac{1}{2n} \text{ und } t_i^k < \frac{1}{2n}.$$

Nun definieren wir $(x_n, y_n, u_n, v_n, t_n) := (x_i^k, y_i^k, u_i^k, v_i^k, t_i^k)$ und mit

$$\|(x_i^k, y_i^k) - (x, y)\| < \frac{1}{n} \text{ und } \|(u_i^k, v_i^k) - (u, v)\| < \frac{1}{n}$$

erhalten wir

$$(x_n, y_n) \xrightarrow{\text{Gph } F} (x, y), (u_n, v_n) \rightarrow (u, v) \text{ und } t_n \searrow 0 \text{ mit } y_n + t_n v_n \in F(x_n + t_n u_n)$$

und somit $v \in TF(x, y; u)$, was zu zeigen war. \square

Die Umkehrung gilt nicht. Dies zeigt eine Betrachtung der Funktion $F(x) = \{x^2, -x^2\}$ im Punkt $(0, 0)$ in Richtung Null abgeleitet oder ebenfalls in $(0, 0)$ die Sägezahnfunktion, die durch $\frac{1}{2}|x|$ beschränkt ist bei beliebiger Richtung. Hier ist der Grenzwert

$$\left[0, \frac{1}{2}|u|\right] \cup \{|u|\} \cup \{-|u|\} \neq [-|u|, |u|].$$

Satz 5.14 ($Cf(x; u)$ vs. $\partial^F f(x)$) Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ um x lokal lipschitzstetig. Dann gilt $Cf(x; u) \neq \emptyset$,

$$Cf(x; u) = \left\{ v = \lim_{t_k \searrow 0} \frac{f(x + t_k u) - f(x)}{t_k} \right\} \text{ und}$$

$$\partial^F f(x) = \{v | \langle v, u \rangle \leq c \forall c \in Cf(x; u) \forall u \in \mathbb{R}^n\}.$$

Im Falle der Richtungs-differenzierbarkeit haben wir somit

$$\partial^F f(x) = \partial^{k_o} f'(x, \cdot)(0).$$

Beweis: Die erste Gleichung ist eine direkte Konsequenz der Definition und der Lipschitzstetigkeit. Die Aussage $Cf(x; u) \neq \emptyset$ folgt auch direkt aus der Lipschitzstetigkeit und der Abschätzung des Differenzenquotienten, denn jede beschränkte Folge besitzt eine konvergente Teilfolge.

Kommen wir also zur Gleichung $\partial^F f(x) = \{v | \langle v, u \rangle \leq c \forall c \in Cf(x; u) \forall u \in \mathbb{R}^n\}$.

„ \subset “: Sei $v \in \partial^F f(x)$, $u \in \mathbb{R}^n$ beliebig und $c \in Cf(x; u)$ mit zugehöriger Folge $t_k \searrow 0$. Der Fall „ $u = 0$ “ ist trivial und wird ausgeschlossen. Nun rechnen wir:

$$\langle v, x + t_k u - x \rangle \leq f(x + t_k u) - f(x) + o(x + t_k u - x) \Rightarrow \left\langle v, \frac{u}{\|u\|} \right\rangle \leq \frac{c}{\|u\|} + \lim_{t_k \searrow 0} \frac{o(t_k u)}{\|t_k u\|},$$

wobei wir bei der Folgerung „ \Rightarrow “ erst durch $t_k \|u\|$ dividiert haben und danach zum Grenzwert übergegangen sind und c eingesetzt haben. Multiplikation mit $\|u\|$ zeigt die gewünschte Inklusion.

„ \supset “: Sei $v \notin \partial^F f(x)$. D.h. $\exists x_k \rightarrow x, \epsilon > 0$ mit $x_k \neq x$ sodass:

$$\left\langle v, \frac{x_k - x}{\|x_k - x\|} \right\rangle \geq \frac{f(x_k) - f(x)}{\|x_k - x\|} + \epsilon.$$

Sei $u_k = \frac{x_k - x}{\|x_k - x\|}$, $t_k = \|x_k - x\|$, $c_k = \frac{f(x_k) - f(x)}{\|x_k - x\|} = \frac{f(x + t_k u_k) - f(x)}{t_k}$. Da $\|u_k\| = 1$ und c_k durch eine lokale Lipschitzkonstante beschränkt ist, können wir zu konvergenten Teilfolgen übergehen und in obiger Ungleichung den Grenzwert betrachten. Also

$$\exists u = \lim_{k \rightarrow \infty} u_k, c := \lim_{k \rightarrow \infty} c_k \in Cf(x; u) : \langle v, u \rangle \geq c + \epsilon.$$

Somit gilt $v \notin \{w | \langle w, u \rangle \leq c \forall c \in Cf(x; u)\}$, was zu zeigen war.

Ist f nun richtungs-differenzierbar, so gilt $Cf(x; u) = \{f'(x; u)\}$ und die letzte Gleichung folgt mit der Definition des konvexen Subdifferentials,

$$\partial^{k_o} g = \{v | \langle v, y - x \rangle \leq t - g(x) \forall t \leq g(y)\}, \text{ also oBdA } t = g(y).$$

□

Folgerung 5.15 (Infimumübergang) Definieren wir zur Funktion f die Funktion

$$f^{lD} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} : f^{lD}(x; u) = \inf \{c \mid c \in Cf(x; u)\},$$

so erhält man analog die Gleichung

$$\partial^F f(x) = \partial^{ko} f^{lD}(x; \cdot)(0).$$

Die verallgemeinerte Ableitung $f^{lD}(x; u)$ wird auch „lower Dini derivative“ genannt, die entsprechende „sup“-Variante „upper Dini derivative“.

Während bei den Normalenkegeln und Subdifferentialen die Mengeneigenschaften bis auf Gleichung 5.1 aus den Definitionen ersichtlich sind, war hier schon etwas mehr Aufwand nötig. Es kommt nun einige Hilfssätze aus [KK02] und letztendlich das Analogon des vorherigen Satzes zur Thibault-Ableitung mit dem Clarke-Subdifferential.

Lemma 5.16 (Zusammenhang) Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine lokal um x lipschitzstetige Funktion. Dann gilt, dass $Cf(x; u)$ und $Tf(x; u)$ zusammenhängende Mengen sind.

Beweisskizze ([KK02]) S.111 Aussage (6.7): Zunächst nimmt man wie üblich an, dass die Aussage falsch ist und wendet die Definition von nicht zusammenhängend, also

$$\exists U_1, U_2 \text{ offen, } U_1 \cap U_2 = \emptyset : M := Tf(x; u) \subset U_1 \cup U_2,$$

an und betrachtet die Differenzenquotienten in einer kleinen Umgebung

$$M(\epsilon) = \left\{ v \mid v = \frac{f(x' + tu) - f(x')}{t}, \|x' - x\| < \epsilon, 0 < t < \epsilon \right\}.$$

Somit hat man $M = \limsup_{\epsilon \searrow 0} M(\epsilon)$ und für ϵ hinreichend klein gilt somit $M(\epsilon) \subset U_1 \cup U_2$. Zu beliebigen $v_{1/2} \in M(\epsilon) \cap U_{1/2}$ mit zugehörigen $x_{1/2}$ und $t_{1/2}$ betrachtet man mit $0 \leq \lambda \leq 1$ die Konvexkombinationen

$$x(\lambda) = \lambda x_1 + (1 - \lambda)x_2 \text{ und } t(\lambda) = \lambda t_1 + (1 - \lambda)t_2$$

$$\text{und sieht sofort } v(\lambda) = \frac{f(x(\lambda) + t(\lambda)u) - f(x(\lambda))}{t(\lambda)} \in M(\epsilon).$$

Somit kann $U_1 \cap U_2 = \emptyset$ nicht gelten, da $v(\lambda)$ beide Mengen stetig verbindet.

Der Beweis für Cf ergibt sich analog durch setzen von $x' = x$. Das Ersetzen von u_k durch u in der Definition von Tf erfolgt wegen der Lipschitzstetigkeit. \square

Lemma 5.17 (Abgeschlossenheit und Beschränktheit) Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine um x lipschitzstetige Funktion mit Rang L . Dann sind $Cf(x; u)$ und $Tf(x; u)$ abgeschlossene Mengen und beschränkt durch $L\|u\|$.

Beweisskizze: Der Beweis erfolgt über eine simple Diagonalisierung und Lipschitzabschätzung der Differenzenquotienten. \square

Folgerung 5.18 (Konvexität) Ist $f : \mathbb{R}^n \rightarrow \mathbb{R}$ um x lokal lipschitzstetig, so sind $Cf(x; u)$ und $Tf(x; u)$ beschränkte abgeschlossene Intervalle und somit auch konvex.

Satz 5.19 ($Tf(x; u)$ vs. $\partial^{Cl}f(x)$) Sei $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ lokal lipschitzstetig um x . Dann gilt

$$\partial^{Cl}F(x)(u) = \text{conv}(TF(x; u)),$$

wobei hier die verallgemeinerte Clarke-Ableitung gemeint ist und wir die einzelnen Elemente auf u anwenden.

Betrachten wir den eindimensionalen Fall $f : \mathbb{R}^n \rightarrow \mathbb{R}$ lokal lipschitz um x , so erhalten wir

$$\partial^{Cl}f(x)(u) = Tf(x; u).$$

Anmerkungen: Die zweite Gleichung folgt direkt aus der ersten Gleichung und der letzten Folgerung.

Der Beweis der ersten Gleichung ist in [KK02] S.115 Theorem 6.6 zu finden.

Die Beweisrichtung „ \subset “ ist sehr elementar zu zeigen, während die andere Richtung den Mittelwertsatz zur verallgemeinerten Clarke-Ableitung braucht. \square

Es folgen zwei Beziehungen des Co-Differentials zu den anderen verallgemeinerten Ableitungen.

Satz 5.20 ($CF(x, y; u)$ vs. Co-Differentiale) Sei $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, dann folgt sofort mit den Definitionen:

$$x^* \in D^F F(x, y)(y^*) \Leftrightarrow \langle x^*, u \rangle \leq \langle y^*, v \rangle \quad \forall u \in \mathbb{R}^n, v \in CF(x, y; u).$$

Der Grenzwertübergang liefert nach Theorem 6.5 in [KK02] für lokal lipschitzstetige Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$:

$$x^* \in D_{lim, F}^F(x)(y^*) \Leftrightarrow \liminf_{x_k \rightarrow x} \sup_{\|u\| \leq 1} (\langle x^*, u \rangle + \sup \langle -y^*, Cf(x; u) \rangle).$$

Für das Co-Differential gibt es einen Hilfssatz. Dieser ist in [RW98] S.366 Satz 9.24 zu finden und besagt:

Lemma 5.21 (Co-Differential vs. Subdifferential) Sei $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine um x lokal lipschitzstetige Funktion. Dann gilt

$$D^\alpha F(x, F(x))(v) = \partial^\alpha (\langle v, F(\cdot) \rangle)(x), \alpha \in \{F; lim, F\}$$

Beweis:

$$\partial^F (\langle v, F(\cdot) \rangle)(x) \stackrel{(4.1)}{=} \{u | \langle u, y - x \rangle \leq \langle v, F(y) - F(x) \rangle + o(y - x)\}.$$

Dies entspricht $(u, -v) \in N_{\text{gph } F}^F(x, F(x))$, da man wie für (4.1) im o -Term den Funktionsanteil durch die Lipschitzstetigkeit „entfernen“ kann.

Die Gleichung für $\alpha = \text{lim}, F$ folgt direkt durch Grenzwertbildung $x_k \rightarrow x, v_k \rightarrow v$, da sich sowohl Normalenkegel als auch Subdifferential ergeben durch Grenzwertbildung aus zu den zugehörigen Mengen des Fréchet-Konzepts. \square

6 Erste Rechenregeln

In diesem Abschnitt wird untersucht, wie sich die benannten Konzepte verhalten, wenn wir zum Beispiel in der zugrundeliegenden Beschreibung mit differenzierbaren Funktionen arbeiten oder die zugrundeliegende Funktion Summe von anderen Funktionen ist. Dies kann einem später helfen Optimalitätsbedingungen zu beweisen oder aber die Berechnungen zu vereinfachen bzw. das Problem der Berechnung auf kleinere Probleme zu reduzieren.

Es gibt natürlich wesentlich mehr als die hier erwähnten Aussagen mit teilweise unterschiedlichen Voraussetzungen. Ich habe mich auf diejenigen beschränkt, die ich für wichtig hielt und vor allem die, die ich später verwende. Zur ausführlichen Lektüre sei hier [RW98] empfohlen.

Die Untersuchung der Ableitung von differenzierbaren Funktionen erfolgt in Abschnitt 10. Dort wird der Normalenkegel untersucht, wenn die betrachtete Menge durch differenzierbare Funktionen beschrieben ist. Mit Satz 8.1 werden wir eine sehr aufwendige Rechenregel für Normalenkegel zeigen.

6.1 ...für Normalenkegel

Die folgende Regel ist zwar recht einfach, aber immer wieder hilfreich.

Satz 6.1 (Kreuzproduktregel) Sei $M = N \times Q \subset \mathbb{R}^n \times \mathbb{R}^q$, dann gilt die Regel

$$N_M^\alpha(x, y) = N_N^\alpha(x) \times N_Q^\alpha(y), \text{ mit } \alpha \in \{ko; F; lim, F; Cl\}$$

Beweisanmerkung: Der Beweis erfolgt durch Fallunterscheidung und durch Betrachten der 2 Teilmengenbeziehungen, wobei für „ \subset “ jeweils auf eine Komponente reduziert wird und für „ \supset “ einfach die zugehörigen Beziehungen addiert werden. Die Ergebnisse für „ lim, F “ und „ Cl “ erhält man sofort durch Grenzwertübergang bzw. Bildung der konvexen Hülle. \square

Im folgenden untersuchen wir, wie sich die Normalenkegel beim Übergang zum Rand von M verhalten.

Satz 6.2 (Normalenkegel an den Rand) Sei $x \in \text{Rand}(M) = M - \text{int}(M)$, dann gilt

$$N_{\text{Rand}(M)}^\alpha(x) \supset N_M^\alpha(x), \alpha \in \{F; lim, F; Cl\}.$$

Beweis: Wir machen eine Fallunterscheidung zu α , wobei „ lF “ = „ lim, F “ gilt.

F : Die Behauptung folgt sofort aus der Definition.

lF : Offensichtlich ist Null in beiden Mengen enthalten. Sei also $v \in N_M^{lim, F}(x)$ mit $v \neq 0$. Wir betrachten Folgen $x_k \xrightarrow{M} x$ und $v_k \xrightarrow{v_k \in N_M^F(x_k)} v$. Es gilt $v_k = 0$, falls $x_k \in \text{int}(M)$. Da $v \neq 0$ können wir oBdA auch $v_k \neq 0$ und somit $x_k \notin \text{int}(M)$ fordern. Somit haben wir $x_k \in \text{Rand}(M)$ und wie eben gezeigt $v_k \in N_{\text{Rand}(M)}^F(x_k)$. Das bedeutet aber $v \in N_{\text{Rand}(M)}^{lim, F}(x)$, wie zu zeigen war.

Cl: Gehen wir nun zur konvexen Hülle über, erhalten wir die Behauptung mit (5.1). □

Gleichheit gilt im Allgemeinen nicht, wie man auch im Abschnitt 3 an den Beispielen am Ende schnell sieht. Hat M allerdings keine inneren Punkte in der Nähe von x , so gilt „lokal“ $\text{Rand}(M) = M$ und wir haben offensichtlich Gleichheit.

Betrachten wir stetige Funktionen f , so gilt $\text{gph } f = \text{Rand}(\text{epi } f)$. Für lokal um x lipschitzstetige Funktionen f haben wir die Formel (4.1):

$$\begin{aligned} \partial^F f(x) &= \{v \mid \langle v, y - x \rangle \leq f(y) - f(x) + o(y - x)\} \\ \Rightarrow \partial^F f(x) &= N_{\text{epi } f}^F(x, f(x)) \cap \mathbb{R}^n \times \{-1\} \\ &= N_{\text{gph } f}^F(x, f(x)) \cap \mathbb{R}^n \times \{-1\}. \end{aligned}$$

Dies liefert nach Multiplikation mit positiven Skalaren die Gleichung:

$$N_{\text{epi } f}^F(x, f(x)) \cap \mathbb{R}^n \times (\mathbb{R}_- - \{0\}) = N_{\text{gph } f}^F(x, f(x)) \cap \mathbb{R}^n \times (\mathbb{R}_- - \{0\}).$$

Betrachten wir nun eine Folge $(x_k, f(x_k)) \rightarrow (x, f(x))$ und zugehörige $(v_k, t_k) \rightarrow (v, t)$ mit $(v_k, t_k) \in N_{\text{gph } f}^F(x_k, f(x_k)) \rightarrow (v, t)$ und $t < 0$. So kann man oBdA $t_k = t < 0$ fordern und erhält die analoge Gleichung auch für „ \lim, F “.

Ein andere Beweis ist in [Mor06] (S.85, Theorem 1.80) zu finden, welcher nur Stetigkeit voraussetzt und die Gleichung

$$D^{\lim, F}(x, f(x))(1) = \partial^{\lim, F} f(x)$$

beweist. Offensichtlich gilt die gezeigte Beziehung nicht für den Clarke-Normalenkegel, wie man schon am Beispiel der Betragsfunktion sieht. Allerdings könnte man in der Gleichung zur konvexen Hülle übergehen.

Eine weitere Formel liefert Lemma 5.21, welches wir für die Berechnung des Co-Differentials von PC^1 -Funktionen verwenden.

6.2 ...für die Co-Differentiale

Die folgende wichtige Summenregel für das Co-Differential ist Theorem 1.62 aus [Mor06], zu finden auf S.70. Wir werden sie später verwenden, um für eine einfache Funktion konkret das Co-Differential zu berechnen.

Satz 6.3 (differenzierbarer Fall und Co-Differential) Sei $g : \mathbb{R}^n \rightarrow \mathbb{R}^q$ im Punkt x differenzierbar und $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^q$ eine mehrwertige Abbildung. Nun gilt

$$D^F(g + G)(x, y)(y^*) = Dg(x)^T y^* + D^F G(x, y - g(x))(y^*).$$

Ist g zusätzlich stetig differenzierbar um x , so folgt die Aussage auch für $D^{\lim, F}(g + G)$.

Beweis: Sei nun also $(x^*, -y^*) \in N_{\text{Gph } g+G}^F(x, y)$, was bedeutet, dass

$$\begin{aligned}
& \langle x^*, x' - x \rangle + \langle -y^*, (g(x') + g') - (g(x) + (y - g(x))) \rangle \\
& \leq o(x' - x) + o(g' - (y - g(x))) \quad \forall g' \in G(x'), x' \in \mathbb{R}^n \\
\Leftrightarrow & \langle x^* - Dg(x)^T y^*, x' - x \rangle + \langle -y^*, g' - (y - g(x)) \rangle \\
& \leq o(x' - x) + o(g' - (y - g(x))) \quad \forall g' \in G(x'), x' \in \mathbb{R}^n \\
\Leftrightarrow & (x^* - Dg(x)^T y^*, -y^*) \in N_{\text{Gph } G}^F(x, y - g(x)).
\end{aligned}$$

Somit ist die erste Gleichung gezeigt.

Durch Grenzwertübergang erhält man wegen „ $\limsup(A + B) \subset \limsup A + \limsup B$ “

$$\begin{aligned}
D^*(g + G)(x, y)(y^*) & \subset Dg(x)^T y^* + D^*G(x, y - g(x))(y^*) \\
& \text{und} \\
D^*((-g) + (g + G))(x, y - g(x))(y^*) & \subset -Dg(x)^T y^* + D^*(g + G)(x, y)(y^*),
\end{aligned}$$

wenn g um x stetig diffbar ist. Die zweite Gleichung erhalten wir durch die Kombination der Beziehungen, weil $D^*G(x, y - g(x))(y^*) = D^*((-g) + (g + G))(x, y - g(x))(y^*)$. \square

Im Beweis in [Mor06] wird auf ein vorheriges Theorem hingewiesen, weshalb ich eine eigene Kurzfassung hier eingefügt habe. Dort wird statt der stetigen Differenzierbarkeit in einer Umgebung strikte Differenzierbarkeit im Punkt x gefordert, was an meinem Beweis nichts ändert.

Definition 6.4 (strikte Differenzierbarkeit) von $g : \mathbb{R}^n \rightarrow \mathbb{R}^q$ im Punkt x

$$\exists Dg(x) \in L(\mathbb{R}^n, \mathbb{R}^q) : \lim_{\substack{\|h\| \searrow 0 \\ x' \rightarrow x}} \frac{\|F(x' + h) - F(x') - DF(x)h\|}{\|h\|} = 0$$

Mit $x' = x$ erhält man die Definition des klassischen Fréchet-Differential. Da dieses eindeutig ist, ist strikte D. eine zusätzliche Forderung an das Verhalten dieser Abbildung bzgl. der Umgebung. Stetige Differenzierbarkeit hat strikte D. zur Folge und ist im Allgemeinen wesentlich leichter nachzuprüfen.

6.3 ...für die Subdifferentiale

Satz 6.5 (separierbare Funktionen) Wir definieren mit $f_i : \mathbb{R}^{n_i} \rightarrow \bar{\mathbb{R}}$ und $n := \sum_{i=1}^m n_i$ eine Funktion $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ durch

$$f(x) := \sum_{i=1}^m f_i(x_i) \text{ mit } x = (x_1, \dots, x_m).$$

Sei f bzw. seien alle f_i in x endlich, dann gilt

$$\partial^\alpha f(x) = \partial^\alpha f_1(x_1) \times \dots \times \partial^\alpha f_m(x_m), \text{ mit } \alpha \in \{ko; F; lim, F; Cl\},$$

wobei die f_i um x_i lokal lipschitzstetig seien für $\alpha = Cl$.

Beweis: Wir machen eine Fallunterscheidung über α , wobei „ lF “=„ lim, F “ gilt:

ko:

$$\begin{aligned}\partial^{ko} f(x) &= \{v | \langle v, y - x \rangle \leq f(y) - f(x)\} \\ &\stackrel{*}{=} \{v | \langle v_i, y_i - x_i \rangle \leq f(y_i) - f(x_i) \forall i\} \\ &= \partial^{ko} f_1(x_1) \times \dots \times \partial^{ko} f_m(x_m)\end{aligned}$$

*Die Inklusion „ \subset “ folgt durch Setzen von $y_j = x_j \forall j \neq i$ mit beliebigem i und die andere „ \supset “ durch Addieren der Ungleichungen.

lF: Dieser Fall wird vollkommen analog bewiesen sogar für nicht-lipschitzstetige Funktionen.

lF: Wenn $f(x_k) = \infty$ so gilt $\partial^F f(x_k) = \emptyset$. Konvergenz im endlichen Kreuzprodukt ist äquivalent zur Konvergenz in den einzelnen Komponenten. Außerdem ergibt sich das Subdifferential aus dem Grenzwert von $\partial^F f(x)$, wie in 4.14 bewiesen.

Cl: Dies folgt aus $\text{conv}(A \times B) = \text{conv}(A) \times \text{conv}(B)$ und Folgerung 5.5.

□

Bemerkung Diese Bedingung ist das Analogon zur Kreuzproduktregel der Normalenkegel.

Satz 6.6 (einfache Beziehungen) Sei $g : X \rightarrow \mathbb{R}$. Aus den Definitionen folgt

$$\begin{aligned}\partial^{ko}(f)(x) + \partial^{ko}(g)(x) &\subset \partial^{ko}(f + g)(x), \\ \partial^F(f)(x) + \partial^F(g)(x) &\subset \partial^F(f + g)(x).\end{aligned}$$

Bemerkung Sucht man bei der Summenbildung nach Gleichungen, so sollte man immer das Beispiel $|x| + (-|x|) = 0$ bedenken. Da im konvexen Fall die Subdifferentialia zusammen fallen, was man auch von jedem Subdifferential erwarten sollte, haben wir $\partial|x|(0) = [-1; 1]$ und $\partial 0 = \{0\}$. Dies zeigt, dass immer zusätzliche Forderungen nötig sind. Für das Subdifferential der konvexen Optimierung braucht man konvexe Funktionen um eine Gleichung zu erhalten. Außerdem werden wir sehen, dass Regularität eine sinnvolle Forderung für die anderen Subdifferentialia ist.

Satz 6.7 (Moreau-Rockafellar-Formel) Seien $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ konvex, dann gilt

$$\partial^{ko}(f)(x) + \partial^{ko}(g)(x) = \partial^{ko}(f + g)(x).$$

Der Beweis erfolgt mittels linearer Transformation und Trennungssatz.

Lemma 6.8 Seien $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ konvex, $f(x) = g(x)$ und $f \leq g$, dann folgt sofort aus der Definition

$$\partial^{ko}(f)(x) \subset \partial^{ko}(g)(x).$$

Satz 6.9 (Clarke-Subdifferential) Für lokal um x lipschitzstetige Funktionen gilt

$$\partial^{Cl}(f)(x) + \partial^{Cl}(g)(x) \supset \partial^{Cl}(f + g)(x).$$

Beweis: Aus der Definition folgt sofort

$$(f + g)^{Cl}(x, u) \leq f^{Cl}(x, u) + g^{Cl}(x, u) \forall u.$$

Mit Lemma 6.8 (denn $0 = (f + g)^{Cl}(x, 0) = g^{Cl}(x, 0) + f^{Cl}(x, 0) = 0 + 0$) und der Darstellung des Clarke-Subdifferentials durch die Richtungsableitung (4.3) erhalten wir

$$\partial^{Cl}(f + g)(x) \subset \partial^{ko}(g^{Cl}(x, u) + f^{Cl}(x, u))(u = 0).$$

Durch die Lipschitzstetigkeit gilt mit Satz 4.15, dass beide Clarke-Ableitungen endlich und konvex sind. Also können wir den Satz von Moreau und Rockafellar anwenden und erhalten die gewünschte Formel. \square

Satz 6.10 (Limiting-Fréchet-Subdifferential)

$$\partial^{lim,F}(f)(x) + \partial^{lim,F}(g)(x) \supset \partial^{lim,F}(f + g)(x). \quad (6.1)$$

Einschränkung und Voraussetzungen: Beim Beweis in [RW98] Korollar 10.9 wird der Satz gezeigt indem man separierbare Funktionen und die Kettenregel aus Theorem 6.14 kombiniert, allerdings erzeugt dies eine aufwendige Regularitätsbedingung. Hier soll reichen, dass f lipschitzstetig ist und g unterhalbstetig, nie $-\infty$ und endlich in x ist, was eine einfache Variante dieser Bedingung ist. \square

Bemerkung Interessanterweise gilt bei $\partial^{lim,F}$ und ∂^{Cl} die umgekehrte Teilmengenbeziehung im Vergleich zu ∂^{ko} und ∂^F . Somit ist $f + g$ regulär, wenn f und g dies sind und wir erhalten eine Summengleichung.

Satz 6.11 (differenzierbarer Fall) Sei f in x differenzierbar, dann gilt

$$\{Df(x)\} + \partial^F(g)(x) = \partial^F(f + g)(x).$$

Ist f weiterhin in einer Umgebung stetig differenzierbar, so gilt auch

$$\begin{aligned} \{Df(x)\} + \partial^{Cl}(g)(x) &= \partial^{Cl}(f + g)(x) \text{ und} \\ \{Df(x)\} + \partial^{lim,F}(g)(x) &= \partial^{lim,F}(f + g)(x). \end{aligned}$$

Beweis: Der Beweis erfolgt analog zu Satz 10.1 ohne Elimination des o -Terms. \square

6.4 ...für die verallgemeinerte Ableitung der Umkehrabbildung

Definition 6.12 (Umkehrabbildung) Sei $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ eine beliebige mehrwertige Abbildung. Dann ist

$$F^{-1} : \mathbb{R}^m \rightrightarrows \mathbb{R}^n, \quad F^{-1}(y) = \{x | y \in F(x)\} \text{ ihre Umkehrabbildung.}$$

Beispiele solcher Abbildungen, wo die Umkehrabbildung relevant ist, sind zum Beispiel das Co-Differential, die Kontingent-Ableitung, die Thibault-Ableitung und die Kojima-Funktion.

Satz 6.13 (Ableitung der Umkehrabbildung) Sei $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ und die zugehörige Umkehrabbildung F^{-1} beliebig. Dann folgt sofort aus den Definitionen

$$\begin{aligned} CF^{-1} : \mathbb{R}^m \rightrightarrows \mathbb{R}^n &= (CF)^{-1}, \\ TF^{-1} : \mathbb{R}^m \rightrightarrows \mathbb{R}^n &= (TF)^{-1} \text{ und} \\ D^\alpha F^{-1} : \mathbb{R}^m \rightrightarrows \mathbb{R}^n &= (D^\alpha F)^{-1}. \end{aligned}$$

Hierbei wurden die verallgemeinerten Ableitungen als mehrwertige Abbildungen verstanden, um die Zusammenhänge sehr einfach notieren zu können.

Ist $f : \mathbb{R} \rightarrow \mathbb{R}$ invertierbar, so haben wir

$$(u, v) \in N_{\text{epi } f}^\alpha(x, f(x)) \Leftrightarrow (v, u) \in N_{\text{epi } f^{-1}}^\alpha(f(x), x).$$

Beweisanmerkungen: Die ersten beiden Beziehungen und die für $\alpha = ko \vee F$ ergeben sich aus den Definitionen. Für $\alpha = \text{"lim, } F\text{"}$ benutzen wir den Grenzwertübergang und $\alpha = Cl$ bilden wir danach die konvexe Hülle und erhalten die Behauptung. \square

6.5 ...für Thibault- und Kontingent-Ableitung

Offensichtlich gilt

Satz 6.14 (positive Skalierbarkeit)

$$CF(x, y; \lambda u) = \lambda CF(x, y; u) \text{ und } TF(x, y; \lambda u) = \lambda TF(x, y; u) \quad \forall \lambda \in \mathbb{R}, \lambda > 0.$$

Für die Summenbildung gibt es die einfachen Regeln:

Satz 6.15 (Summenformel) Seien $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^m$, lokal lipschitzstetig um x , dann gilt

$$\begin{aligned} T(f + g)(x; u) &\subset Tf(x; u) + Tg(x; u), \\ C(f + g)(x; u) &\subset Cf(x; u) + Cg(x; u) \text{ und} \\ T(f)(x; u + v) &\subset Tf(x; u) + Tf(x; v). \end{aligned}$$

Beweisskizze: Im wesentlichen braucht man nur die Definitionen und die Umformungen

$$\begin{aligned}
 & (f(x_k + t_k u) + g(x_k + t_k u)) - (f(x_k) + g(x_k)) \\
 = & (f(x_k + t_k u) - f(x_k)) + (g(x_k + t_k u) - g(x_k)), \\
 & (f(x + t_k u) + g(x + t_k u)) - (f(x) + g(x)) \\
 = & (f(x + t_k u) - f(x)) + (g(x + t_k u) - g(x)) \text{ und} \\
 & f(x_k + t_k(u + v)) - f(x_k) \\
 = & (f((x_k + t_k v) + t_k u) - f(x_k + t_k v)) + (f(x_k + t_k v) - f(x_k)).
 \end{aligned}$$

Division durch t_k und Verwendung konvergenter Teilfolgen, die wegen der Lipschitzstetigkeit existieren, liefert die gewünschten Teilmengenbeziehungen nach Grenzwertbildung. \square

Bemerkung Für die Kontingent-Ableitung kann man keine Aussage bzgl. der Summe der Ableitungsrichtungen machen. Dies sieht man am „Schnapper“-Beispiel 7. Mit $u = (1, 1)$ und $v = -u$, erhält man

$$Cf(0, u + v) = Cf(0, 0) = \{0\} \text{ und } Cf(0, u) = Cf(0, v) = \{1\}.$$

Man sieht auch, dass man für die Thibault-Ableitung eine echte Inklusion hat:

$$Tf(0, 0) = \{0\} \subset [-2; 2] = Tf(0, u) + Tf(0, v).$$

Auch bei der Summe von Funktionen erhalten wir für beide Ableitungen eine echte Inklusion. Ein einfaches Beispiel zur Thibault-Ableitung ist $f = |x|, g = -f$. Für die Kontingent-Ableitung braucht man ein komplexeres Beispiel mit mehrwertiger Kontingent-Ableitung. Zum Beispiel kann man dann statt $|x|$ die Sägezahnfunktion zu $\frac{1}{2}|x|$ nehmen. Gleichheit erhält man für die Kontingent-Ableitung, wenn eine der Funktionen richtungsdifferenzierbar ist mittels $(-f)'(x; u) = -(f'(x; u))$ und $g = (g+f) + (-f)$. Für die Thibault-Ableitung gilt Gleichheit, wenn f um x stetig differenzierbar ist.

Ein „Gegenbeispiel“ mit einer nur stetigen Funktion wäre $f = |x|^{-3}$ und $g = -f$:

$$T(f + g)(0; u) = \{0\}, \quad Tf(0, u) = -Tg(0, u) = \emptyset \vee \mathbb{R} \text{ und}$$

$$Tf(0; 0) = \mathbb{R} \not\subset \emptyset = Tf(0; 1) + Tf(0; -1).$$

Nun wollen wir das Verhalten der Ableitungen bei der Verkettung von Funktionen untersuchen. Zunächst betrachten wir den einfachen Fall

$$\begin{aligned}
 g &: \mathbb{R}^n \rightarrow \mathbb{R}^m \text{ lokal lipschitzstetig,} \\
 h &: \mathbb{R}^m \rightarrow \mathbb{R}^p \text{ lokal lipschitzstetig,} \\
 f &: \mathbb{R}^n \rightarrow \mathbb{R}^p, f := h(g(\cdot)), y := g(x).
 \end{aligned}$$

Hierzu gibt es in [KK02] Abschnitt 6.4.1 eine Kettenregel.

Satz 6.16 (einfache Kettenregel)

$$\begin{aligned}
 Tf(x; u) & \subset \bigcup_{v \in Tg(x; u)} Th(g(x); v) =: Th(g(x); Tg(x; u)), \\
 Cf(x; u) & \subset \bigcup_{v \in Cg(x; u)} Ch(g(x); v) =: Ch(g(x); Cg(x; u)).
 \end{aligned}$$

Ist g richtungsdifferenzierbar, so haben wir Gleichheit bei der Kontingent-Ableitung. Für die Thibault-Ableitung erhalten wir Gleichheit, wenn g stetig differenzierbar ist um x .

Beweisskizze: Der Beweis folgt direkt aus den Definitionen, da man wegen der Lipschitzstetigkeit von h zu konvergenten Teilfolgen übergehen kann und wegen der zusätzlichen Lipschitzstetigkeit von g die Ableitungen nicht leer sind.

$$v_k := \frac{g(x_k + t_k u) - g(x_k)}{t_k} \Rightarrow \frac{h(g(x_k + t_k u)) - h(g(x_k))}{t_k} = \frac{h(g(x_k) + t_k v_k) - h(g(x_k))}{t_k}$$

Ist g richtungsdifferenzierbar, so enthält die Kontingent-Ableitung nur die Richtungsableitung. Sei L eine lokale Lipschitzkonstante zu h bzgl. $g(x)$. Mit

$$\begin{aligned} & \left\| \frac{h(g(x + t_k u)) - h(g(x))}{t_k} - \frac{h(g(x) + t_k g'(x; u)) - h(g(x))}{t_k} \right\| \\ &= \left\| \frac{h(g(x + t_k u)) - h(g(x) + t_k g'(x; u))}{t_k} \right\| \leq L \left\| \frac{g(x + t_k u) - g(x)}{t_k} - g'(x, u) \right\| \end{aligned}$$

folgt dann die vorletzte Aussage. Die letzte erhält man, indem man x_k statt x verwendet und wegen $g'(x, u) = Dg(x)u$. \square

Beispiel 13 (Gegenbeispiel zur Thibault-Ableitung) Dass die erste Inklusion eine echte sein kann, sieht man wieder an der Schnapperfunktion Beispiel 7, $h : \mathbb{R}^2 \rightarrow \mathbb{R}$, $h(x, y) = \min \{|x|, |y|\}$ schon mit der linearen Funktion $g : \mathbb{R}^2 \rightarrow \mathbb{R}$, $g(x, y) = (x, 0)$, denn

$$\begin{aligned} f(x, y) &\equiv 0 \\ Tf((x, y); u) &= \{0\}, Tg((x, y); u) = \{(u_1, 0)\}, \\ Th((0, 0); (u_1, 0)) &= [-|u_1|, |u_1|] = Th(g(0, y); v) \text{ mit } v \in Tg((0, y); u), \\ Tf((x, y); u) &\neq Th((0, 0); (u_1, 0)) \quad \forall u : u_1 \neq 0. \end{aligned}$$

Beispiel 14 (Gegenbeispiel zur Kontingent-Ableitung) Für die Kontingent-Ableitung braucht man ein komplizierteres Beispiel, da g nicht richtungsdifferenzierbar sein darf.

Sei $h : \mathbb{R} \rightarrow \mathbb{R}$ die Sägezahnfunktion zu $\frac{1}{2}|x|$, siehe Beispiel 5. Wir setzen $g = h$ und betrachten $f(x) = h(h(x))$. Man rechnet dann aus

$$Cf(0; u) = \left[0, \frac{1}{5}|u|\right] \neq Ch(0; Ch(0; u)) = Ch\left(0; \left[0, \frac{1}{2}|u|\right]\right) = \left[0, \frac{1}{4}|u|\right].$$

Berechnungsskizze: Da die Gleichung $Cf(0; u) = \left[0, \frac{1}{5}|u|\right]$ nicht offensichtlich ist, folgt eine zugehörige Beweisskizze. Aufgrund der Symmetrie, genügt es den Fall $u \geq 0$ zu betrachten und wegen der Lipschitzstetigkeit und positiven Skalierbarkeit untersuchen wir nur der Fall $u = 1$.

Da wir wissen, dass die Kontingent-Ableitung zusammenhängend ist, brauchen wir nur die untere und obere Grenze zu ermitteln. Die untere ist offensichtlich 0. Die obere ist kleiner als $\frac{1}{4}$.

Zunächst untersuchen wir h , die Sägezahnfunktion. Eine Betrachtung der entsprechenden Dreiecke zeigt, dass die Nullstellen bzw. lokalen Minima bei $y_k = \frac{1}{3^k}$ liegen und die Schnitte mit $\frac{1}{2}|x|$ bzw. lokalen Maxima bei $z_k = \frac{2}{3^k}$. Es gilt $h(z_k) = \frac{1}{3^k} = y_k$ und somit $f(z_k) = f(y_k) = 0$. Wegen des Anstieges 1 bzw. -1 von h auf den Intervallen $\left[\frac{1}{3^k}, \frac{2}{3^k}\right]$ bzw. $\left[\frac{2}{3^k}, \frac{3}{3^k}\right]$ gilt $f(x) = h(|x - \frac{1}{3^k}|) \forall x \in \left[\frac{1}{3^k}, \frac{3}{3^k}\right]$.

Für die Kontingent-Ableitung im Punkt 0 in Richtung 1 müssen wir die Werte $\frac{f(t)}{t}$ betrachten, da $f(0) = 0$. Man sieht schnell ein, dass diese maximal werden im Intervall $\left[\frac{1}{3^k}, \frac{3}{3^k}\right]$ im Punkt

$$t_k = \frac{5}{3^{k+1}} = \frac{2}{3^k} - \frac{1}{3^{k+1}} \text{ mit } f(t_k) = \frac{1}{3^{k+1}}.$$

Dort haben wir $\frac{f(t_k)}{t_k} = \frac{1}{5}$ auch im Grenzwert $k \rightarrow \infty$. □

Beispiel 15 (Gegenbeispiel zur Lipschitzstetigkeit) Warum g und h lokal lipschitzstetig sein sollten ist klar, da sonst die entsprechenden Differentiale leer sein können. Ein Beispiel wäre $g(x) = 0$, $h(y) = \text{sign}(y)|y|^{-3}$ bzw. $h(y) = 0$, $g(x) = \text{sign}(x)|x|^{-3}$.

Satz 6.17 (differenzierbarer Fall) Fordern wir nun zusätzlich, dass h um x stetig differenzierbar ist, so erhalten wir Gleichheit.

Beweisskizze: Zunächst macht man eine Fallunterscheidung und betrachtet den Fall $g(x_k + t_k u) - g(x_k) = 0$ extra. Dann erhält man

$$\frac{h(g(x_k + t_k u)) - h(g(x_k))}{t_k} = Dh(x_k) \frac{g(x_k + t_k u) - g(x_k)}{t_k} + \frac{o(g(x_k + t_k u) - g(x_k))}{t_k}.$$

Wegen der lokalen Lipschitzstetigkeit von g erhalten wir

$$o(g(x_k + t_k u) - g(x_k)) = o(t_k).$$

Für die Inklusion „ \subset “ gehen wir über zu einer konvergenten Teilfolge bzgl. g . Und für die Inklusion „ \supset “ können wir die Produktregel bei existierenden Grenzwerten anwenden, da Dh in x stetig ist.

Der Beweis für die Kontingent-Ableitung erfolgt mittels $x_k \equiv x$. □

Da die Forderung der Differenzierbarkeit eine recht starke Forderung ist, die man nicht immer erfüllen kann, betrachtet man einen Spezialfall der Verkettung.

$$\begin{aligned} g : \mathbb{R}^m &\rightarrow \mathbb{R}^q && \text{lokal lipschitzstetig,} \\ h : \mathbb{R}^{n+q} &\rightarrow \mathbb{R}^p && \text{lok. lip. und } D_g h \text{ existiere und sei auch lok.lip.,} \\ f : \mathbb{R}^{n+m} &\rightarrow \mathbb{R}^p, && f(x, y) := h(x, g(y)). \end{aligned}$$

Der Hintergedanke dabei ist wahrscheinlich, dass man versucht, ungünstige Komponenten der Funktion zu trennen, um sie „unabhängig“ voneinander zu betrachten. Die Variablen der zweiten Komponente von h werden intuitiv mit g benannt.

Für diese Struktur erhalten wir

Satz 6.18 (spezielle Kettenregel) Seien f, g, h wie eben beschrieben. Dann gilt

$$Tf((x, y); (u, v)) \subset T_x h((x, g(y)); u) + T_g h((x, g(y)); Tg(x; v)), \quad (6.2)$$

$$Cf((x, y); (u, v)) \subset C_x h((x, g(y)); u) + C_g h((x, g(y)); Cg(x; v)) \quad (6.3)$$

mit $T_g h(x, g(y); v') = C_g h(x, g(y); v') = D_g h(x, g(y))v'$.

Beweisskizze: Der vorige Satz liefert nur

$$\begin{aligned} Tf((x, y); (u, v)) &\subset Th((x, g(y)); (u, Tg(x; v))), \\ Cf((x, y); (u, v)) &\subset Ch((x, g(y)); (u, Cg(x; v))), \text{ denn} \\ (u, Tg(x; v)) &= T(x', g(y'))((x, g(y)); (u, v)) \end{aligned}$$

und analog für $C(x', g(y'))$, da man das u aus den entsprechenden Grenzwert ziehen kann. Sei also $v' \in Tg(x; v)$. Wegen der Lipschitzstetigkeit können wir in den Definitionen u und v' betrachten statt u_k und v'_k . Seien also $x_k \rightarrow x$, $g_k \rightarrow g(y)$ und $t_k \searrow 0$ entsprechende Folgen aus der Definition für die Mengen der rechten Seite der Teilmengenbeziehung bzgl. (u, v') . Sei L eine lokale Lipschitzkonstante zu h und $D_g h$ um $(x, g(y))$. Der Rest folgt durch zweckmäßiges Aufspalten und Abschätzen:

$$\begin{aligned} &\frac{h(x_k + t_k u, g_k + t_k v') - h(x_k, g_k)}{t_k} \\ = &\frac{h(x_k + t_k u, g_k + t_k v') - h(x_k + t_k u, g_k)}{t_k} + \frac{h(x_k + t_k u, g_k) - h(x_k, g_k)}{t_k} \\ = &D_g h(x_k + t_k u, g_k)v' + \frac{o(t_k)}{t_k} + \frac{h(x_k + t_k u, g_k) - h(x_k, g_k)}{t_k} \text{ und} \\ &\left\| \frac{h(x_k + t_k u, g_k) - h(x_k, g_k)}{t_k} - \frac{h(x_k + t_k u, g(y)) - h(x_k, g(y))}{t_k} \right\| \\ = &\left\| \frac{h(x_k + t_k u, g_k) - h(x_k + t_k u, g(y))}{t_k} - \frac{h(x_k, g_k) - h(x_k, g(y))}{t_k} \right\| \\ = &\left\| \int_0^1 D_g h(x_k + t_k u, g(y) + sw)sw - D_g h(x_k, g(y) + sw)sw \, ds \right\| \frac{1}{t_k} \\ &\text{mit } w = (g_k - g(y)) \\ = &\left\| \int_0^1 (D_g h(x_k + t_k u, g(y) + sw) - D_g h(x_k, g(y) + sw))sw \, ds \right\| \frac{1}{t_k} \\ \leq &\int_0^1 s \|D_g h(x_k + t_k u, g(y) + sw) - D_g h(x_k, g(y) + sw)\| \cdot \|w\| \, ds \frac{1}{t_k} \\ \leq &L \|t_k\| \|u\| \|g_k - g(y)\| \frac{1}{t_k} = L \|u\| \|g_k - g(y)\| \rightarrow 0. \end{aligned}$$

Eine ausführlichere Rechnung ist unter anderem in [KK02] Theorem 6.8 zu finden. \square

Leider erhalten wir keine Gleichung. Um nun die dazu nötige „unabhängige“ Betrachtung der Grenzwerte machen zu können, brauchen wir eine zusätzliche Forderung an g . Diese wird in [KK02] *simple (einfach)* genannt und ist unterschiedlich je nachdem, welche Ableitung man betrachtet. Im wesentlichen fordert man, dass die Wahl der t_k in den jeweiligen Definitionen für g keine Rolle spielt.

Definition 6.19 (einfach bzgl. Thibault-Ableitung) Sei $g : \mathbb{R}^m \rightarrow \mathbb{R}^q$ lokal lipschitzstetig, dann heißt g einfach in $y \in \mathbb{R}^m$, wenn

$$\forall v \in \mathbb{R}^m, w \in Tg(y; v), t_k \searrow 0, \exists y_k \rightarrow y : \frac{g(y_k + t_k v) - g(y_k)}{t_k} \rightarrow w$$

zumindest für eine Teilfolge von $k \rightarrow \infty$.

Einfache Funktionen sind zum Beispiel differenzierbare Funktionen, $g(y) = y$, $g(y) = y^+$ mit $y_i^+ = \max\{y_i, 0\}$ und $g(y) = (y^+; y - y^+) =: (y^+; y^-)$.

Definition 6.20 (einfach bzgl. Kontingent-Ableitung) Sei $g : \mathbb{R}^m \rightarrow \mathbb{R}^q$ lokal lipschitzstetig, dann heie g einfach in $y \in \mathbb{R}^m$, wenn

$$\forall v \in \mathbb{R}^m, w \in Cg(y; v), t_k \searrow 0 : \frac{g(y + t_k v) - g(y)}{t_k} \rightarrow w$$

zumindest für eine Teilfolge von $k \rightarrow \infty$. Mit anderen Worten: g ist richtungsdifferenzierbar und somit $Cg(y; v) = \{g'(y; v)\}$.

Die entsprechende Formulierung für mehrwertige Abbildungen wäre „proto-differentiability“ und wird in [RW98] definiert und untersucht. Das Analogon für die Thibault-Ableitung gibt es dort nicht.

Seien nun f, g, h wie zuvor beschrieben und g zusätzlich entsprechend *einfach*, dann ist in [KK02] Theorem 6.8 folgender Satz bewiesen.

Satz 6.21 (spezielle Kettenregelgleichung)

$$Tf((x, y); (u, v)) = T_x h((x, g(y)); u) + D_g h(x, g(y)) Tg(x; v), \quad (6.4)$$

$$Cf((x, y); (u, v)) = C_x h((x, g(y)); u) + D_g h(x, g(y)) Cg(x; v) \quad (6.5)$$

Beweisskizze: Die Beziehung „ \subset “ haben wir im letzten Satz bewiesen, ohne dass g *einfach* ist. Für die Rückrichtung „ \supset “ wird diese Eigenschaft dann benutzt.

Sei also (u, v) beliebig. Wir betrachten zugehörige $x_k \rightarrow x$, $t_k \searrow 0$ für ein Element aus $T_x h((x, g(y)); u)$. Weil g *einfach* ist, erhalten wir zu einem Element aus $Tg(x; v)$ und zu unserer Folge $x_k \rightarrow x$ passende $y_k \rightarrow y$, wie in der Definition. Der zweite neue Trick außer dieser geschickten Folgenwahl besteht darin, die Stetigkeit der Ableitung auszunutzen:

$$\begin{aligned} \lim_{\substack{x'_k \rightarrow x \\ v_k \rightarrow v}} D_g h(x'_k) v_k &= \lim_{\substack{x'_k \rightarrow x \\ v_k \rightarrow v}} (D_g h(x'_k) v_k - D_g h(x) v_k) + (D_g h(x) v_k - D_g h(x) v + D_g h(x) v) \\ &= 0 + 0 + D_g h(x) v = D_g h(x) v \end{aligned}$$

Der Rest ist dann eine ähnliche Rechnung, wie im vorigen Beweis. Um das ganze notieren zu können, definieren wir: $x'_k = x_k + t_k u$ und $y'_k = y_k + t_k v$.

$$\begin{aligned}
& \frac{h(x'_k, g(y)) - h(x_k, g(y))}{t_k} + D_g h(x'_k) \frac{g(y'_k) - g(y_k)}{t_k} \\
= & \frac{h(x'_k, g(y)) - h(x_k, g(y))}{t_k} - \frac{h(x'_k, g(y_k)) - h(x'_k, g(y'_k))}{t_k} + \frac{o(t_k)}{t_k} \\
= & \frac{h(x'_k, g(y'_k)) - h(x_k, g(y_k))}{t_k} + \frac{o(t_k)}{t_k} \\
- & \left\{ \frac{h(x'_k, g(y_k)) - h(x_k, g(y_k))}{t_k} - \frac{h(x'_k, g(y)) - h(x_k, g(y))}{t_k} \right\}.
\end{aligned}$$

Nun können wir zu konvergenten Teilfolgen übergehen wegen der lokalen Lipschitzstetigkeit von g und h . Außerdem schätzen wir wieder ab:

$$\begin{aligned}
& \left\| \frac{h(x_k + t_k u, g(y_k)) - h(x_k, g(y_k))}{t_k} - \frac{h(x_k + t_k u, g(y)) - h(x_k, g(y))}{t_k} \right\| \\
\leq & L \|u\| \|g(y_k) - g(y)\| \rightarrow 0, \text{ wie im letzten Beweis, mit } g_k = g(y_k).
\end{aligned}$$

Die Teilmengenbeziehung ergibt sich dann durch Grenzwertübergang in der Gleichung. Für die Kontingent-Ableitung kann man wieder $x_k \equiv x$ setzen. \square

Nun schließt sich gleich eine Anwendung des Satzes für das Produkt von Matrizen an.

Folgerung 6.22 (Produktregel) Seien $M : \mathbb{R}^n \rightarrow \mathbb{R}^{a \times b}$ und $N : \mathbb{R}^q \rightarrow \mathbb{R}^{b \times c}$ lokal lipschitzstetige matrixwertige Funktionen und eine von beiden sei entsprechend der berechneten Ableitung *einfach*. Für $F : \mathbb{R}^{n+q} \rightarrow \mathbb{R}^{a \times c}$, $F(x, y) := M(x)N(y)$ erhalten wir dann

$$TF((x, y); (u, v)) = TM(x; u)N(y) + M(x)TN(y; v), \quad (6.6)$$

$$CF((x, y); (u, v)) = CM(x; u)N(y) + M(x)CN(y; v) \quad (6.7)$$

Beweis: Wenn N einfach ist, setzen wir $g = N$ und $h(x, z) = M(x)z$. $N(y)$ lässt sich dann aus dem entsprechenden Grenzwert von $T_x h((x, g(y)); u)$ rausziehen und $D_g h(x, g(y)) = M(x)$. Alle anderen Beweise folgen analog. \square

Dieser Satz aus [KK02] Korollar 6.10 wird verwendet um die Ableitung der Kojima-Funktion auszurechnen, deren Nullstellen im Allgemeinen optimalverdächtige Punkte sind.

7 Stabilität - Co-Differential - Regularität

In diesem Kapitel stellen wir den Zusammenhang mit einer Art Injektivitätsbedingung des Co-Differentials und einer Stabilitätsbedingung bzw. Stetigkeitsbedingung der mehrwertigen Abbildung her, welche Aubin property (im folgenden Aubin-Eigenschaft) genannt wird.

Danach zeigen wir die Äquivalenz dieser Bedingungen zu einer Regularitätsbedingung, die im späteren Verlauf noch öfter auftauchen wird. Ziel dieser Umformulierung ist es die Regularitätsbedingung besser zu verstehen, mit anderen Bedingungen zu vergleichen und letztendlich uns mit der Anwendung bei der konkreten Berechnung zu beschäftigen.

7.1 Aubin-Eigenschaft

Definition 7.1 (Aubin-Eigenschaft) Eine mehrwertige Abbildung $F : P \rightrightarrows X$ besitzt die Aubin-Eigenschaft in $(\bar{p}, \bar{x}) \in \text{Gph } F$, wenn es $\epsilon, \delta, L > 0$ gibt, sodass:

$$F(p) \cap B(\bar{x}, \epsilon) \subset F(p') + B(0, L \|p - p'\|) \quad \forall p, p' \in B(\bar{p}, \delta).$$

Eine andere verbreitete Bezeichnung ist pseudo Lipschitzstetigkeit. Alternativ betrachtet man auch F^{-1} und dann heißt die Eigenschaft metrisch regulär oder pseudo-regulär.

Mit $p = \bar{p}$ sieht man sofort wegen $\bar{x} \in F(\bar{p})$, dass $F(p') \neq \emptyset \quad \forall p'$ nahe \bar{p} .

Statt ϵ und δ kann man zum Minimum übergehen und nur eine der beiden Größen verwenden ohne die Definition zu ändern.

Damit wir mit dieser Eigenschaft besser umgehen können, wollen wir diese zunächst umformulieren. Dazu gehen wir zur Negation über:

Lemma 7.2 (Negation) F hat nicht die Aubin-Eigenschaft mit Lipschitzkonstante L in (\bar{p}, \bar{x}) genau dann, wenn

$$\begin{aligned} \exists (p_k, x_k) \rightarrow (\bar{p}, \bar{x}) \text{ und } p'_k \rightarrow \bar{p}, \text{ sodass } (p_k, x_k) \in \text{Gph } F \text{ und} \\ \text{dist}(x_k, F(p'_k)) > L \|p'_k - p_k\| > 0 \quad \forall k \in \mathbb{N} \end{aligned} \quad (7.1)$$

Der folgende Satz aus [Kum08] Lemma 11.1.3 formuliert diese Bedingung sinnvoll um.

Satz 7.3 (Anwendung Ekeland's Variationsprinzips auf die negierte Aubin-Eigenschaft) Sei F abgeschlossen. Aus (7.1) folgt mit $\lambda := L$:

$$\begin{aligned} \exists (p_k^E, x_k^E) \xrightarrow{\text{Gph } F} (\bar{p}, \bar{x}) \text{ und } p'_k \rightarrow \bar{p}, \text{ mit } p_k^E \neq p'_k, \text{ sodass} \\ (p_k^E, x_k^E) \text{ die Funktion } H_\lambda(p, x) := \|p - p'_k\| + \frac{1}{\lambda} d(x, x_k^E) \text{ auf } \text{Gph } F \text{ minimiert.} \end{aligned} \quad (7.2)$$

Die Umkehrung folgt für $L \in (0, \lambda)$.

Beweis: Die letzte Aussage zeigt man sehr einfach.

„ \Leftarrow “ Dazu setzen wir $(p_k, x_k) := (p_k^E, x_k^E)$ und erhalten mit $p := p'_k$ und der Minimaleigenschaft:

$$\frac{1}{\lambda} d(x, x_k) \geq \|p_k - p'_k\| \quad \forall x \in F(p'_k).$$

„ \Rightarrow “ Gelte (7.1). Wir zeigen nun damit (7.2). Mit

$$\begin{aligned} h(x) := \text{dist}(p'_k, F^{-1}(x)), \quad \epsilon_k := \|p'_k - p_k\| \text{ und } (p_k, x_k) \in \text{Gph } F \text{ folgt} \\ 0 \leq \inf_{x \in \mathbb{R}^n} h(x) \leq h(x_k) = \text{dist}(p'_k, F^{-1}(x_k)) \leq 0 + \epsilon_k. \end{aligned}$$

Nun können wir mit $\alpha_k := L\epsilon_k$ Ekeland's Variationsprinzip anwenden und erhalten

$$x_k^E \in B(x_k, \alpha_k) \text{ mit } h(x) + \frac{1}{L}d(x, x_k^E) \geq h(x_k^E) \quad \forall x \in \mathbb{R}^n \text{ und } h(x_k^E) \leq h(x_k).$$

Folglich gilt $F^{-1}(x_k^E) \neq \emptyset$ und da F abgeschlossen ist, wählen wir

$$p_k^E \in F^{-1}(x_k^E) \text{ mit } \|p_k^E - p_k'\| = \min_{p \in F^{-1}(x_k^E)} \|p - p_k'\|.$$

Somit minimiert (p_k^E, x_k^E) die Funktion $H_\lambda(p, x)$ auf $\text{Gph } F$, denn

$$H_\lambda(p_k^E, x_k^E) = \|p_k^E - p_k'\| = h(x_k^E) \leq h(x) + \frac{1}{\lambda}d(x, x_k^E) \leq H(x, p) \text{ mit } p \in F^{-1}(x).$$

Mit $p_k', p_k \rightarrow \bar{p}, x_k \rightarrow \bar{x}, x_k^E \in B(x_k, \alpha_k)$ und $\alpha_k = L\|p_k' - p_k\|$ folgt $(p_k^E, x_k^E) \rightarrow (\bar{p}, \bar{x})$.

Bleibt $p_k^E \neq p_k'$ zu zeigen. Dies folgt über Widerspruch mit

$$\alpha_k \geq \|x_k - x_k^E\| \geq \text{dist}(x_k, F(p_k^E)) \stackrel{p_k^E = p_k'}{=} \text{dist}(x_k, F(p_k')) \stackrel{(7.1)}{>} L\|p_k' - p_k\| = \alpha_k.$$

□

Erfüllt nun F die Aubin-Eigenschaft nicht, so existieren die o.g. Folgen für beliebige L , also haben wir, da wir L beliebig wählen können, die „Äquivalenz“ der beiden oben genannten Formulierungen. Diese verwenden wir nun für die Rückübersetzung der Negation für eine alternative Beschreibung der Aubin-Eigenschaft.

Folgerung 7.4 (Aubin-Eigenschaft nach Anwendung von Ekeland's Variationsprinzip) $F : P \rightrightarrows X$ besitzt die Aubin Eigenschaft in $(\bar{p}, \bar{x}) \in \text{Gph } F$ genau dann, wenn es $\epsilon, \delta, L > 0$ gibt, sodass:

$$\begin{aligned} &\forall p, p' \in B(\bar{p}, \epsilon), x \in F(p) \text{ mit } p \neq p' \text{ gilt:} \\ &\exists p_e \in P : L\|p_e - p'\| + \text{dist}(x, F(p_e)) < L\|p - p'\| \end{aligned}$$

Beweisanmerkung: p_e steht hier für das Element, dass Optimalität des obigen $H_L(q, y)$ verhindert in $(p, x) = (p_k^E, x_k^E)$ nahe (\bar{p}, \bar{x}) mit $p' = p_k'$. □

Beliebt ist auch die folgende Charakterisierung der Aubin-Eigenschaft in [Mor06] S.48 Theorem 1.41, welche aber für unsere Beweise nicht benötigt wird.

Satz 7.5 (Charakterisierung über Abstandsfunktion) $F : P \rightrightarrows X$ besitzt die Aubin-Eigenschaft in $(\bar{p}, \bar{x}) \in \text{Gph } F$ genau dann, wenn $g : P \times X \rightarrow \mathbb{R}, g(p, x) = \text{dist}(x, F(p))$ die Aubin-Eigenschaft in (\bar{p}, \bar{x}) besitzt.

7.2 Limiting Epsilon-Normalen und das Co-Differential

In diesem Abschnitt zeigen wir eine etwas andere „Konstruktion“ von Elementen des Limiting-Fréchet-Normalenkegels. Diese wenden wir dann auf das Co-Differential an und erhalten für spezielle Richtungen eine bessere Charakterisierung davon. Zunächst definieren wir neue Normalenkegel.

Definition 7.6 (ϵ -Normalenkegel) Zunächst definieren wir den ϵ -Normalenkegel

$$N_M^{\alpha, \epsilon}(x) = \{v \mid \langle v, y - x \rangle \leq \epsilon \|y - x\| \forall y \in M \text{ nahe } x\}.$$

Die Elemente von $N_M^{\alpha, \epsilon}(x)$ werden auch ϵ -Normalen genannt.

Nun können wir, wie beim Limiting-Fréchet-Normalenkegel, zum Grenzwert übergehen, wobei wir zusätzlich die Epsilon gegen Null gehen lassen und erhalten den Limiting- ϵ -Normalenkegel

$$N_M^{lim, \epsilon}(x) = \limsup_{x_k \xrightarrow{M} x, \epsilon_k \searrow 0} N_M^{\alpha, \epsilon_k}(x_k).$$

Bemerkung (Alternative Definition) Wir haben

$$|\langle v_k, y - x_k \rangle - \langle v, y - x_k \rangle| \leq \|v_k - v\| \|y - x_k\| \forall y \in M \text{ nahe } x.$$

Somit können wir $v_k \equiv v$ wählen, was die Definition vereinfacht. Dazu addieren wir zum ursprünglichen ϵ_k die Norm $\|v_k - v\|$ hinzu und erhalten

$$N_M^{lim, \epsilon}(x) = \left\{ v \mid \exists x_k \xrightarrow{M} x, \epsilon_k \searrow 0, \langle v, y - x_k \rangle \leq \epsilon_k \|y - x_k\| \forall y \in M \text{ nahe } x_k \right\}.$$

Aufgrund der Ähnlichkeit stellt sich die Frage nach dem Zusammenhang zum Limiting-Fréchet-Normalenkegel und der folgende Satz aus [Kum08] Theorem 9.1.4 zeigt, dass es keinen Unterschied gibt.

Satz 7.7 ($N_M^{lim, \epsilon}(x) = N_M^{lim, F}(x)$)

Beweis:

$$N_M^F(x_k) \subset N_M^{\alpha, \epsilon}(x_k) \forall \epsilon > 0 \Rightarrow N_M^{lim, F}(x) \subset N_M^{lim, \epsilon}(x)$$

Für die Umkehrung betrachten wir nun $v \in N_M^{lim, \epsilon}(x)$ mit zugehörigen $x_k \xrightarrow{M} x$ und $\epsilon_k \searrow 0$ aus der vereinfachten Formel, d.h. $v \in N_M^{\alpha, \epsilon_k}(x_k)$. OBdA gelte $\|v\| = 1$ und $\epsilon_k < \frac{1}{2}$. Nun betrachten wir $x_k(t) = x_k + tv$ und die zugehörige euklidische Projektion auf die Menge M : $\pi_k(t)$, wobei $\pi_k(t)$ hinreichend „nahe“ x_k sei, wie in der Definition von $N_M^{lim, \epsilon}(x)$. $t > 0$ muss nur hinreichend klein gewählt werden, da $x_k(t)$ und $\pi_k(t)$ gegen x_k konvergieren mit $t \searrow 0$. Nun haben wir die Beziehungen:

$$\begin{aligned} t &= \langle v, x_k(t) - x_k \rangle \text{ und} \\ t &= \|x_k(t) - x_k\| \geq \|x_k(t) - \pi_k(t)\|, \quad \text{da } x_k \in M \text{ und folgern} \\ 2t &\geq \|x_k(t) - \pi_k(t)\| + \|x_k(t) - x_k\| \geq \|\pi_k(t) - x_k\| \quad \text{und somit} \\ 2\epsilon_k t &\geq \langle v, \pi_k(t) - x_k \rangle, \quad \text{da } v \in N_M^{lim, \epsilon}(x) \text{ und } t \text{ hinreichend klein und} \\ t &> 2\epsilon_k t \geq \langle v, \pi_k(t) - x_k \rangle = \langle v, x_k(t) - x_k \rangle = t, \quad \text{falls zusätzlich } x_k(t) = \pi_k(t). \end{aligned}$$

Somit folgt $x_k(t) \neq \pi_k(t)$ und wegen der Konstruktion von $\pi_k(t)$ über euklidische Projektion gilt $u_k(t) := \frac{x_k(t) - \pi_k(t)}{\|x_k(t) - \pi_k(t)\|} \in N_M^F(\pi_k(t))$. Sei $t_k \searrow 0$ eine beliebige Folge. Durch Übergang zu einer Teilfolge gilt $u_k(t_k) \rightarrow u$, da $\|u_k(t_k)\| = 1 \forall k$. Also gilt $u \in N_M^{lim,F}(x)$. Mithilfe der bisherigen Berechnungen zeigen wir, dass Konvergenz gegen v vorliegt.

$$\begin{aligned}
\|v - u_k(t_k)\|^2 &= \|v\|^2 + \|u_k(t_k)\|^2 - 2\langle v, u_k(t_k) \rangle \\
&= 2 - 2\left\langle v, \frac{x_k(t_k) - \pi_k(t_k)}{\|x_k(t_k) - \pi_k(t_k)\|} \right\rangle \\
&\leq 2 - 2\frac{1}{t_k} \langle v, x_k(t_k) - \pi_k(t_k) \rangle \\
&= 2 - 2\frac{1}{t_k} (\langle v, x_k(t_k) - x_k \rangle + \langle v, x_k - \pi_k(t_k) \rangle) \\
&\leq 2 + 2\frac{1}{t_k} (-t_k + 2\epsilon_k t_k) \\
&= 4\epsilon_k \searrow 0
\end{aligned}$$

Also gilt $v = u \in N_M^{lim,F}(x)$. Somit haben wir auch $N_M^{lim,F}(x) \supset N_M^{lim,\epsilon}(x)$ gezeigt. \square

Folgerung 7.8 (Anwendung auf $\text{Gph } F$) Speziell für $M := \text{Gph } F$ bedeutet dies nun:

$$u \in D^*F(x, y)(-v) \Leftrightarrow (u, v) \in N_{\text{Gph } F}^{lim,F}(x, y) \Leftrightarrow$$

$$\exists \epsilon_k \searrow 0, z_k \xrightarrow{\text{Gph } F} (x, y) : (u, v) \in N_{\text{Gph } F}^{\alpha, \epsilon_k}(z_k).$$

(u, v) ist also eine ϵ_k -Normale in $z_k = (x_k, y_k)$, d.h.

$$\langle u, x' - x_k \rangle + \langle v, y' - y_k \rangle \leq \epsilon_k (\|x' - x_k\| + \|y' - y_k\|) \forall (x', y') \in \text{Gph } F \text{ nahe } z_k.$$

Bemerkung (Nullrichtungen) Nun interessieren wir uns für $u \in D^*F(x, y)(0)$ mit $u \neq 0$. Wir setzen also $v = 0$ und können oBdA $\|u\| = 1$ fordern, da $N_{\text{Gph } F}^{lim,F}(x, y)$ ein Kegel ist. Wollen wir nun, dass solche u nicht existieren, so negieren wir die Aussage und erhalten folgenden einfachen Satz.

Satz 7.9 (Co-Differential und Nullrichtungen) $D^*F(x, y)(0) = \{0\} \Leftrightarrow$

$$\exists \epsilon > 0 : \forall z = (x^z, y^z) \in \text{Gph } F, \|z - (x, y)\| < \epsilon \nexists (u, 0) \in N_{\text{Gph } F}^{\alpha, \epsilon}(z) \text{ mit } \|u\| = 1.$$

Mit dieser Charakterisierung stellen wir nun im folgenden Abschnitt den Bezug zur Aubin-Eigenschaft her.

7.3 Mordukhovich-Kriterium

Wir haben bereits mehrere Beschreibungen der Aubin-Eigenschaft und des Co-Differentials. Jetzt folgt der Zusammenhang der Stabilitätsbedingung mit der „Injektivität des Co-Differentials“.

Satz 7.10 (Mordukhovich-Kriterium) Sei $F : X \rightrightarrows Y$ eine mehrwertige Abbildung und der Graph sei in $(x, y) \in \text{Gph } F$ lokal abgeschlossen. Dann hat F die Aubin-Eigenschaft (ist pseudo lipschitz) in (x, y) genau dann, wenn

$$D^*F(x, y)(0) = \{0\}.$$

Widerspruchsbeweise: Wir betrachten die zwei Beweisrichtungen.

(“ \Rightarrow “) Sei F also pseudo lipschitz in (x, y) . Angenommen $\exists u \in D^*F(x, y)(0)$ mit $u \neq 0$, dann erhalten wir mit Folgerung 7.8 und anschließender Bemerkung

$$\begin{aligned} & \exists \epsilon_k \searrow 0, (x_k, y_k) \xrightarrow{\text{Gph } F} (x, y) : (u, 0) \in N_{\text{Gph } F}^{\alpha, \epsilon_k}(x_k, y_k), \text{ d.h.} \\ & \langle u, x' - x_k \rangle \leq \epsilon_k (\|x' - x_k\| + \|y' - y_k\|) \quad \forall (x', y') \in \text{Gph } F \text{ nahe } (x_k, y_k). \end{aligned}$$

Nun haben wir, dass F pseudo lipschitz ist, was bedeutet,

$$\exists \epsilon, L : F(\dot{x}) \cap B(y, \epsilon) \subset F(x') + B(0, L \|\dot{x} - x'\|) \quad \forall \dot{x}, x' \in B(x, \epsilon).$$

OBdA gelte $\|x_k - x\| < \frac{\epsilon}{2}$ und mit $t < \frac{\epsilon}{2}$ folgt $\|(x_k + tu) - x\| \leq \|x_k - x\| + \|tu\| \leq 2\frac{\epsilon}{2}$. Es gelte weiterhin $\|y_k - y\| < \epsilon$ und einsetzen von $\dot{x} := x_k$ und $x' := x_k + tu$ liefert:

$$y_k \in F(x_k + tv) + B(0, Lt), \text{ d.h. } \exists y_k(t) \in F(x_k + tv) : \|y_k - y_k(t)\| \leq Lt.$$

Nun wählen wir ein spezielles t_k so klein, dass $((x', y') := (x_k + t_k v, y_k(t_k)))$ „nahe“ (x_k, y_k) ist und erhalten mit der ersten Ungleichung

$$t_k = \langle u, x' - x_k \rangle \leq \epsilon_k (\|x' - x_k\| + \|y' - y_k\|) \leq \epsilon_k (t_k + Lt_k).$$

Nach Division durch t_k erhalten wir $1 \leq \epsilon_k(1 + L)$ und somit einen Widerspruch mittels $\epsilon_k \searrow 0$. Folglich gilt $u = 0$.

(“ \Leftarrow “) Sei F also nicht pseudo lipschitz in (x, y) . Mit (7.2) folgt dann

$$\begin{aligned} \forall \epsilon > 0 \exists (x_k, y_k) \xrightarrow{\text{Gph } F} (x, y), x_k \neq x'_k \rightarrow x : (x_k, y_k) \text{ minimiert} \\ H_\epsilon(x', y') = \|x' - x'_k\| + \epsilon \|y' - y_k\| \text{ auf } \text{Gph } F. \end{aligned}$$

Also kann man $\epsilon_k \searrow 0$ statt ϵ wählen und erhält dieselbe Bedingung. Jetzt werden wir die Minimumsaussage umformen und eine spezielle Umgebung wählen um (x_k, y_k) für die (x', y') . Wir definieren $v_k = x'_k - x_k$ und die Minimumsaussage liefert

$$0 < \|v_k\| \leq \|x' - x'_k\| + \epsilon_k \|y' - y_k\| = \|(x' - x_k) - v_k\| + \epsilon_k \|y' - y_k\|.$$

Nun können wir die Ungleichung quadrieren und erhalten mittels Parallelogrammgleichung und Dreiecksungleichung

$$0 < \|x' - x_k\|^2 - 2 \langle v_k, x' - x_k \rangle + 2\epsilon_k (\|x' - x_k\| + \|v_k\|) \|y' - y_k\| + \epsilon_k^2 \|y' - y_k\|^2.$$

Fordern wir für unsere $(x', y') \in \text{Gph } F$ zusätzlich, sie sollen nahe (x_k, y_k) sein, genauer $\|x' - x_k\| < \epsilon_k \|v_k\|$ und $\|y' - y_k\| \leq \epsilon_k \|v_k\|$, so folgt aus der Ungleichung

$$2 \langle v_k, x' - x_k \rangle < \epsilon_k \|v_k\| \|x' - x_k\| + (2\epsilon_k^2 + 2\epsilon_k + \epsilon_k^3) \|v_k\| \|y' - y_k\|.$$

Nun folgt mit $u_k := \frac{v_k}{\|v_k\|}$ und $\bar{\epsilon}_k := \epsilon_k^2 + \epsilon_k + \epsilon_k^3 \searrow 0$

$$\langle u_k, x' - x_k \rangle < \bar{\epsilon}_k (\|x' - x_k\| + \|y' - y_k\|) \quad \forall (x', y') \in \text{Gph } F \text{ nahe } (x_k, y_k).$$

Also $u_k \in N_{\text{Gph } F}^{\alpha, \bar{\epsilon}_k}(x_k, y_k)$ mit $\|u_k\| = 1$. Gehen wir nun zur konvergenten Teilfolge über so erhalten wir, wie im vorhergehenden Abschnitt gezeigt ein $u \in D^*F(x, y)(0)$ mit $\|u\| = 1$ und somit haben wir einen Widerspruch, da $D^*F(x, y)(0) \neq \{0\}$. \square

Der Beweis ist auch in [RW98] zu finden (S.380, Theorem 9.40) und ist dort recht aufwendig, wir haben hier diesen Beweis an [KK02] angelehnt, der zuvor beide Bedingungen umformuliert. Der Beweis ist hier vergleichsweise einfach, da wir nur den endlichdimensionalen Fall behandeln. Im unendlichen Fall erhält man keine Äquivalenz-Beziehung. Allgemeinere Aussagen befinden sich in [KK02] ab S.66 Theorem 3.7.

Nun betrachten wir $S := F^{-1} : Y \rightrightarrows X$. Hierbei sind die Urbilder gemeint: $S(y) = \{x | y \in F(x)\}$. Die zugehörigen Graphen sind gleich bis auf eine „Achsenvertauschung“. Somit gilt die Relation:

$$(u, v) \in N_{\text{Gph } F}^{\text{lim}, F}(x, y) \Leftrightarrow (v, u) \in N_{\text{Gph } F^{-1}}^{\text{lim}, F}(y, x)$$

und ein Einsetzen in die Definition liefert

$$D^*F^{-1}(y, x)(u) = \left\{ v \in Y \mid (-u, v) \in N_{\text{Gph } F}^{\text{lim}, F}(x, y) \right\}.$$

$D^*F(x, y)(0) = \{0\}$ bedeutet also nichts anderes als die Implikation:

$$(u, 0) \in N_{\text{Gph } F}^{\text{lim}, F}(x, y) \Rightarrow u = 0$$

Mit obiger Umkehrabbildungseigenschaft liefert dies:

$$0 \in D^*F^{-1}(x, y)(u) \Rightarrow u = 0.$$

Jetzt erkennt man erst richtig, wieso zu Anfang von einer Injektivitätsbedingung geredet wurde, denn hier tritt eine Verallgemeinerung dieser Bedingung aus der linearen Algebra auf.

7.4 Limiting-Fréchet-Regularitätsbedingung bzw. GMFCQ

Wir definieren nun, mit den im Abschnitt 4.3 eingeführten Größen g , N und Q , die mehrwertige Abbildung $F : \mathbb{R}^q \rightrightarrows \mathbb{R}^n$:

$$F(p) := \{x \in N \mid g(x) + p \in Q\}.$$

Die zugehörige Umkehrabbildung hat eine etwas schönere Gestalt, welche im folgenden Satz hergeleitet und verwendet wird.

Es gibt zu den Größen eine Regularitätsbedingung, welche den Limiting-Fréchet-Normalenkegel benutzt. Sie wird auch GMFCQ genannt, abgeleitet von verallgemeinerte MFCQ. Den Zusammenhang werden wir am Ende des nächsten Kapitels zeigen.

Definition 7.11 (GMFCQ) für $M := N \cap g^{-1}(Q)$ in $x \in M$

$$\left. \begin{array}{l} Dg(x)^T y^* \in -N_N^{\text{lim}, F}(x) \\ y^* \in N_Q^{\text{lim}, F}(g(x)) \end{array} \right\} \Rightarrow y^* = 0, \quad (7.3)$$

Satz 7.12 (GMFCQ vs. Aubin) GMFCQ in $x \in M$ ist genau dann erfüllt, wenn F die Aubin-Eigenschaft hat in $(0, x)$ bzw. $D^*F(0, x)(0) = \{0\}$

Beweis: In [Out06] S.9 Theorem 1.1 ist ein ähnlicher Satz zu finden. Allerdings das Nachrechnen wird dem Leser überlassen. Die dazu nötigen Rechenregeln haben wir bereits bewiesen und werden sie nun anwenden. Nach dem Mordukhovich-Kriterium genügt es $D^*F(0, x)(0) = \{0\}$ mit GMFCQ zu vergleichen. Ich wähle lieber die zuvor erwähnte äquivalente Bedingung über die Umkehrabbildung $S := F^{-1}$

$$0 \in D^*S(x, 0)(u) \Rightarrow u = 0.$$

Diese Bedingung hat schon im entferntesten Sinne die Gestalt der GMFCQ. Ein großer Vorteil ist die einfache Gestalt von

$$S(x) = \begin{cases} Q - g(x), & \text{falls } x \in N \\ \emptyset, & \text{sonst.} \end{cases}$$

Definieren wir uns nun die Hilfsabbildung $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^q$ mittels:

$$G(x) = \begin{cases} Q, & \text{falls } x \in N \\ \emptyset, & \text{sonst.} \end{cases},$$

so erhalten wir $S(x) = -g(x) + G(x)$ und können die Rechenregel aus Satz 6.3 anwenden, da g stetig differenzierbar ist.

$$D^*(S)(x, y)(y^*) = -Dg(x)^T y^* + D^*G(x, y + g(x))(y^*).$$

Mit $\text{Gph } G = N \times Q$ und Satz 6.1 erhalten wir

$$D^*(S)(x, y)(y^*) = -Dg(x)^T y^* + N_N^{\text{lim}, F}(x), \text{ falls } -y^* \in N_Q^{\text{lim}, F}(y + g(x))$$

und ansonsten die leere Menge. Somit heißt $0 \in D^*S(x, 0)(y^*) \Rightarrow y^* = 0$ nichts anderes als

$$\left. \begin{array}{l} -y^* \in N_Q^{\text{lim}, F}(g(x)) \\ Dg(x)^T y^* \in N_N^{\text{lim}, F}(x) \end{array} \right\} \Rightarrow y^* = 0.$$

Dies entspricht eindeutig der GMFCQ, da man die Mengenbeziehungen natürlich mit (-1) multiplizieren kann. \square

8 Anwendung der GMFCQ - Hauptrechenregel

Der folgende Satz ist die Grundlage aller Optimalitätsbedingungen von J. Outrata [Out06], mit denen ich mich beschäftigen werde. Er liefert bei Regularität eine Kettenregel bzw. Rechenregel für die Normalenkegel.

Satz 8.1 nach R.T. Rockafellar und R. Wets 1998 [RW98]

Sei $g : \mathbb{R}^n \rightarrow \mathbb{R}^q$ stetig differenzierbar, $N \subset \mathbb{R}^n$ und $Q \subset \mathbb{R}^q$ abgeschlossene Mengen, und $x \in M := N \cap g^{-1}(Q)$.

Dann gilt

$$N_M^F(x) \supset Dg(x)^T N_Q^F(g(x)) + N_N^F(x).$$

Unter der zusätzlichen Regularitätsbedingung (GMFCQ)

$$\left. \begin{array}{l} Dg(x)^T u^* \in -N_N^{lim,F}(x) \\ u^* \in N_Q^{lim,F}(g(x)) \end{array} \right\} \Rightarrow u^* = 0,$$

erhalten wir sogar

$$N_M^F(x) \subset Dg(x)^T N_Q^{lim,F}(g(x)) + N_N^{lim,F}(x)$$

in einer kleinen Umgebung, da dort Regularität gilt. Durch Abgeschlossenheit von $N_H^{lim,F}$ und somit der rechten Seite zeigt man

$$N_M^{lim,F}(x) \subset Dg(x)^T N_Q^{lim,F}(g(x)) + N_N^{lim,F}(x).$$

Die Aussage M ist in x normal regulär, wenn dies N in x und Q in $g(x)$ sind, folgt sofort.

Beweisanmerkungen: Der Beweis steht in [RW98] unter Theorem 6.14, Seite 208. Auch in dem Buch spielt der Satz eine größere Rolle.

Der Beweis der ersten Teilmengenbeziehung ist einfach zu zeigen, wie in Abschnitt 4.3 bereits geschehen.

Für die „umgekehrte“ Teilmengenbeziehung wird die Charakterisierung eines Normalenvektors als Ableitung einer Funktion in ihrem Minimalpunkt auf der betrachteten Menge benutzt. Dann wendet man eine Strafmethode an und erhält im Grenzwert die gewünschte Aussage. \square

Bemerkung Die GMFCQ gilt durch die Grenzwertbildung auch in einer kleinen Umgebung. Außerdem reduziert sie sich auf den Fréchet-Kegel, wenn N und Q normal regulär sind.

Folgerung 8.2 (Clarke-Konzept) Die obigen Sätze gelten auch mit dem Clarke-Normalenkegel statt dem Limiting-Fréchet-Normalenkegel mittels Gleichung (5.1) und Folgerung 4.7, denn lineare Abbildungen kann man vorziehen beim Bilden der konvexen Hülle. Da der Fréchet-Normalenkegel auch konvex ist, bleiben die Aussagen bzgl. Gleichungen mit Limiting-Fréchet-Normalenkegel die selben bei Ersetzung durch den Clarke-Normalenkegel. Die Regularitätsbedingung mit Clarke-Normalenkegel statt der Limiting-Version zu verwenden, wäre eine stärkere Forderung, aber ist im Allgemeinen einfacher zu prüfen.

Im folgenden gehe ich auf wichtige Spezialfälle des Satzes ein.

Folgerung 8.3 Setzen wir $N = \mathbb{R}^n$, so bedeutet die Regularitätsbedingung, dass $Dg(x)^T$ injektiv ist auf $N_Q^{lim,F}(g(x))$. Mit

$$N_M^{lim,F}(x) \subset Dg(x)^T N_Q^{lim,F}(g(x))$$

erhalten wir die umgekehrte Beziehung zum Fréchet-Fall. Zu beachten ist, dass die Injektivität auch im allgemeinen Fall gefordert wird.

Dieser Fall wird später nochmal ausführlicher diskutiert. Genauso gehe ich dann noch auf spezielle Q ein. Es wird sich dann auch der Bezug der Regularitätsbedingung GMFCQ zur Aubin-Eigenschaft und MFCQ zeigen.

Folgerung 8.4 (Normalenkegel bei Mengenschnitt) Sei $x \in M := N \cap Q$, dann gilt

$$N_M^F(x) \supset N_Q^F(x) + N_N^F(x).$$

Ist zusätzlich die Regularitätsbedingung $N_Q^{lim,F}(x) \cap -N_N^{lim,F}(x) = \{0\}$ bzw. $u \in N_Q^{lim,F}(x), v \in N_N^{lim,F}(x) : u + v = 0 \Rightarrow u, v = 0$ erfüllt, so gilt außerdem

$$N_M^{lim,F}(x) \subset N_Q^{lim,F}(x) + N_N^{lim,F}(x),$$

also wieder die Umkehrung zum Fréchet-Fall. Übergang zur konvexen Hülle liefert:

$$N_M^{Cl}(x) \subset N_Q^{Cl}(x) + N_N^{Cl}(x).$$

Gleichheit gilt, wenn N und Q zusätzlich normal regulär sind in x .

Beweis: Es genügt $g = id$ in Satz 8.1 einzusetzen. Die beiden Regularitätsbedingungen sagen genau dasselbe aus. Für den Clarke-Kegel braucht man wieder Gleichung (5.1) und Folgerung 4.7 \square

Folgerung 8.5 (Normalenkegel bei Gleichungsrestriktion) Sei $M = g^{-1}(\{0\})$, dann gilt

$$N_M^F(x) \supset Dg(x)^T \mathbb{R}^q.$$

Sind die Gradienten Dg_i^T zusätzlich linear unabhängig, so gilt

$$N_M^\alpha(x) = Dg(x)^T \mathbb{R}^q \quad \forall \alpha \in \{F; lim, F; Cl\}.$$

Beweis: $N = \mathbb{R}^n$ und $Q = \{0\}$ sind normal regulär, da sie konvex sind, und die lineare Unabhängigkeit der Gradienten entspricht der GMFCQ. Die Gleichung für den Clarke-normalenkegel folgt wieder über die konvexe Hülle und Gleichung (5.1). \square

Satz 8.6 (äquivalenter Satz ohne N) Setzen wir $N := \mathbb{R}^n$ ein, so lautet die GMFCQ

$$\left. \begin{array}{l} Dg'(x)^T u^* = 0 \\ u^* \in N_{Q'}^{lim,F}(g'(x)) \end{array} \right\} \Rightarrow u^* = 0$$

und wir haben die beiden Teilmengenbeziehungen

$$N_M^F(x) \supset Dg'(x)^T N_{Q'}^F(g'(x)) \quad \text{und} \quad N_M^{lim,F}(x) \subset Dg'(x)^T N_{Q'}^{lim,F}(g'(x))$$

mit $M = g'^{-1}(Q')$.

Setzen wir in diesem „neuen“ Satz nun $Q' = N \times Q$ und $g' : \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^q$, $g'(x) = (x, g(x))$ ein, so erhalten wir wieder den alten Satz. Die entsprechende Aussage für Gleichheit bei normaler regulären Mengen reduziert sich nun auf Q' und für Äquivalenz braucht man den einfachen Satz, dass ein Kreuzprodukt von Mengen genau dann normal regulär ist, wenn die beiden Mengen normal regulär sind.

Beweis:

$$(u^*, v^*)^T \in N_{Q'}^{lim, F}(g'(x)) \stackrel{(6.1)}{\Leftrightarrow} u^* \in N_N^{lim, F}(x), v^* \in N_Q^{lim, F}(g(x))$$

$$Dg'(x)^T (u^*, v^*)^T = 0 \Leftrightarrow u^* = -Dg^T v^*$$

Beide Bedingungen zusammen liefern genau die alte GMFCQ.

Die Teilmengenbeziehungen aus dem alten Satz ergeben sich durch einsetzen der Gleichungen $Dg'(x)^T = (\mathbb{E}_n, Dg^T)$ und $N_{Q'}^\alpha(g'(x)) \stackrel{(6.1)}{=} N_N^\alpha(x) \times N_Q^\alpha(g(x))$ in die neue Teilmengenbeziehung. \square

8.1 Vergleich der Anwendung der GMFCQ mit der klassischen MFCQ

Im folgenden werde ich auf einen Spezialfall unserer Hauptrechenregel eingehen, in der die GMFCQ mit einer anderen bekannten Regularitätsbedingung zusammenfällt, der Mangasarian-Fromovitz-Constraint-Qualification (kurz: MFCQ).

Definition 8.7 (MFCQ) Seien $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^q$ stetig differenzierbar und

$$M = \{x | h(x) = 0, g(x) \leq 0\}.$$

Dann erfüllt $x \in M$ die MFCQ genau dann, wenn die $Dh_i(x)$ linear unabhängig sind und es eine Richtung $u_0 \in \mathbb{R}^n$ gibt mit $Dh(x)u_0 = 0$ und $Dg_j(x)u_0 < 0 \forall j : g_j(x) = 0$. Diese wird auch MFCQ-Richtung genannt.

Bemerkung Diese Definition ist zwar die übliche für MFCQ, allerdings werden manchmal auch andere Bedingungen so oder als MFCQ-artige Bedingungen bezeichnet.

Da u_0 positiv skalierbar ist, fordert man manchmal

$$g(x) + Dg(x)u_0 \in \text{int}(\mathbb{R}_+^q)$$

statt der letzten Bedingung.

Satz 8.8 (MFCQ vs. GMFCQ) Sei $Q' = 0^p \times \mathbb{R}_+^q$ und $g' : \mathbb{R}^n \rightarrow \mathbb{R}^{p+q}$ stetig differenzierbar mit komponentenweiser Darstellung

$$g' = (h, g) \text{ und } x \in M = g'^{-1}(Q) = \{x | h(x) = 0, g(x) \leq 0\}.$$

Nun sind GMFCQ und MFCQ in x äquivalent und der Satz 8.1 liefert unter der Regularitätsbedingung die bereits bekannte Formel

$$N_M^\alpha(x) = \left\{ Dh(x)^T u + Dg(x)^T v \mid u \in \mathbb{R}^p, v \in \mathbb{R}_+^q, v_j g_j(x) = 0 \forall j \right\} \forall \alpha \in \{F; lim, F; Cl\}.$$

Beweis: Zunächst berechnen wir

$$N_{Q'}^F(g'(x)) = N_{Q'}^{\text{lim},F}(g'(x)) \stackrel{(6.1)}{=} \mathbb{R}^p \times \prod_{j=1}^q \begin{cases} 0, & \text{falls } g_j(x) < 0 \\ [0, \infty), & \text{falls } g_j(x) = 0 \end{cases}$$

und $Dg'(x)^T = (Dh(x)^T, Dg(x)^T)$. Da in beiden Bedingungen die j mit $g_j(x) < 0$ nicht relevant sind, gehen wir, damit die Notation einfacher ist, oBdA davon aus, dass alle g_j aktiv sind, d.h. $g(x) = 0$. GMFCQ bedeutet also in diesem Fall:

$$Dh(x)^T u + Dg(x)^T v = 0, (u, v) \in \mathbb{R}^p \times \mathbb{R}_+^q \Rightarrow (u, v) = 0.$$

(„GMFCQ \Rightarrow MFCQ“) Setzen wir nun zunächst $v = 0$ in die GMFCQ ein, so bedeutet sie die lineare Unabhängigkeit der $Dh_i(x)$, wie sie auch in der MFCQ gefordert ist. Bilden wir jetzt das Skalarprodukt einer MFCQ-Richtung u_0 mit der GMFCQ-Gleichung ergibt dies

$$(Dh(x)u_0)^T u + (Dg(x)u_0)^T v = (Dg(x)u_0)^T v = 0.$$

Da $v \geq 0$ und $Dg(x)u_0 < 0$ ergibt dies sofort $v = 0$. Mit der bereits gezeigten linearen Unabhängigkeit der $Dh_i(x)$ folgt dann $u = 0$. Folglich impliziert die MFCQ die GMFCQ.

(„MFCQ \Rightarrow GMFCQ“) Umgekehrt sei nun GMFCQ erfüllt und es ist noch eine MFCQ-Richtung u_0 zu finden. Betrachten wir zunächst die Hilfsmenge

$$H := \left\{ (u, v) \in \mathbb{R}^p \times \mathbb{R}_+^q \mid \sum_{j=1}^q v_j = 1 \right\}.$$

Man sieht sofort, dass H konvex und abgeschlossen ist. Da $Dg'(x)^T : \mathbb{R}^p \times \mathbb{R}^q \rightarrow \mathbb{R}^n$ eine lineare Abbildung ist, erhalten wir zunächst, dass $Dg'(x)^T(H)$ eine konvexe, abgeschlossene Menge ist, welche nicht Null enthält. Da $\{0\}$ auch eine konvexe, abgeschlossene Menge ist, können wir beide Mengen echt voneinander trennen und erhalten ein u_0 mit

$$(Dh(x)u_0)^T u + (Dg(x)u_0)^T v < 0 \quad \forall (u, v) \in H.$$

Sei zunächst $u = 0$, so erhalten wir die von der MFCQ-Richtung geforderte Bedingung $Dg_j(x)u_0 < 0$ indem wir für v den j -ten Einheitsvektor einsetzen. Gilt $Dh_i(x)u_0 \neq 0$ so wählen wir zu beliebigen zulässigen v unser u als Vielfaches des i -ten Einheitsvektors mit $t(Dh_i(x)u_0)$ und der Grenzwert in der Ungleichung mit $t \rightarrow \infty$ liefert einen Widerspruch. Somit haben wir auch gezeigt, dass u_0 die andere Bedingung: $Dh(x)u_0 = 0$ erfüllt und folglich eine MFCQ-Richtung ist.

(Normalenkegelgleichung) Der Beweis der Gleichung folgt aus der Formel für den Normalenkegel von Q' an $g'(x)$, weil Q' normal regulär ist, und Satz 8.1. \square

Eine andere Beweisidee ist in [RW98] S.226 Beispiel 6.40 zu finden. Diese ist jedoch umfangreicher, da sie den Umweg über die Tangentialkegel nutzt.

9 Optimalitätsbedingungen

Bisher haben wir ausführlich die verallgemeinerten Ableitungen untersucht. Auch wurde an Beispielen untersucht, was für Aussagen diese Ableitungen über die Funktion haben können. In diesem Abschnitt wird nun die für die Optimierung wichtigste Eigenschaft untersucht. Wie erkennt man optimale (minimale) bzw. optimalverdächtige Punkte? Wir werden sehen, dass die wichtigen Bedingungen einfach sind und die Schwierigkeit wirklich im Ermitteln der Ableitungen besteht. Am Ende des Abschnitts werden wir noch gesondert die Probleme andeuten, wie wenig eine Optimalitätsbedingung aussagen kann über das Minimum, selbst wenn man mit differenzierbarer Zielfunktion arbeitet.

Aus der konvexen Optimierung gibt es den Satz:

Satz 9.1 (konvexe Optimierung) x ist ein globales Minimum der Zielfunktion f genau dann, wenn

$$0 \in \partial^{ko} f(x)$$

gilt. Ist die Restriktionsmenge M zusätzlich konvex und abgeschlossen und f konvex, so ist

$$0 \in \partial^{ko} f(x) + N_M^{ko}(x)$$

eine notwendige und hinreichende Bedingung für globale Optimalität auf M .

Dieser Satz wird für die Verallgemeinerungen in ähnlichen Versionen auftauchen. Leider sind die folgenden Optimalitätsbedingungen alle nur notwendig. Es gibt auch hinreichende Bedingungen, die aber nicht notwendig sind. Diese werden in dieser Arbeit aber nicht betrachtet.

Satz 9.2 (Fréchet-Konzept) Seien $f : X \rightarrow \mathbb{R}$ und $M \subset X$ gegeben und x sei ein lokal minimaler Punkt von f auf M , der abgeschlossenen Restriktionsmenge. Unter speziellen weiteren Voraussetzungen, die im folgenden aufgelistet sind, gelten dann die folgenden notwendigen Optimalitätsbedingungen in x :

1. Fall

Vor.: Sei f differenzierbar in x

Opt.Bed.: $-Df(x) \in N_M^F(x)$ bzw. $0 \in Df(x) + N_M^F(x)$

Bew.: $Df(x)(y - x) = f(y) - f(x) + o(y - x) \leq o(y - x) \forall y \in M$

Bem.: $Df(x)u \geq 0 \forall u \in T_M^B(x)$ ist die „polare“ Bedingung.

2. Fall

Vor.: Gelte $M = X$. Für diesen Satz ist $f(y) = \infty$ möglich außer für $y = x$.

Opt.Bed.: $0 \in \partial^F f(x)$

Bew.: $\langle 0, y - x \rangle = 0 \leq f(y) - f(x) + o(y - x) \forall y \in X$

3. Fall

Vor.: $\partial_M^F f(x) \stackrel{Def.}{=} \{x^* | \langle x^*, y - x \rangle \leq f(y) - f(x) \forall y \in M\}$

Opt.Bed.: $0 \in \partial_M^F f(x)$

Bew.: siehe 2.Fall

Bem.: Diese Bedingung ist eher unbrauchbar, da sie im Allgemeinen sehr schwer zu berechnen wäre. Außerdem ist eine getrennte Berechnung von Größen zur Menge und zur Funktion besser.

4. Fall

Vor.: $\partial^F(-f)(x) \neq \emptyset$

Opt.Bed.: $0 \in -\partial^F(-f)(x) + N_M^F(x)$ bzw. $\partial^F(-f)(x) \subset N_M^F(x)$

Bew.: $\partial^F(-f)(x) \subset \partial_M^F(-f)(x) \subset N_M^F(x)$

Bem.: Schon für einfache Funktionen wie $|x|$ ist die Voraussetzung verletzt.

Bem.: Diese Art der Bedingung wäre höchstens brauchbar für den Clarke-Normalenkegel, da hier $-\partial^{Cl}(-f)(x) = \partial^{Cl}(f)(x)$ gilt, wenn f lokal lipschitzstetig um x ist. Der Spezialfall $M = X$ zeigt jedoch, dass mit dieser Bedingung nichts sinnvolles zu erwarten ist.

Bem.: Auch wenn diese Bedingung abstrus erscheint, ist sie im [RW98] relevant, wobei dort eine Extradefinition eingeführt wird, für das entsprechende Subdifferential.

5. Fall

Vor.: Sei f richtungsdifferenzierbar auf $T_M^B(x)$ und in x lokal lipschitzstetig.

Opt.Bed.: $f'(x; u) \geq 0 \forall u \in T_M^B(x)$

Bew.: $\exists u_k \rightarrow u, t_k \searrow 0 : x + t_k u_k \in M$. L sei eine lokale Lipschitzkonstante.

$$\begin{aligned} \left\| \frac{f(x + t_k u_k) - f(x)}{t_k} \right\| &\leq L \|u_k\| \Rightarrow \text{oBdA Konvergenz} \\ \frac{f(x + t_k u_k) - f(x)}{t_k} &\geq 0 \\ \left\| \frac{f(x + t_k u_k) - f(x)}{t_k} - \frac{f(x + t_k u) - f(x)}{t_k} \right\| &\leq L \|u_k - u\| \end{aligned}$$

Somit nimmt $f'(x; u)$ den selben Grenzwert „ ≥ 0 “ an.

6. Fall

Vor.: Sei f in x lokal lipschitzstetig.

Opt.Bed.: $\forall u \in T_M^B(x), \exists v \in Cf(x; u) : v \geq 0$

Bew.: klar, da entsprechendes v aus der Konstruktion von u resultiert, siehe 5.Fall.

In den folgenden 2 Sätzen wird die Teilmengenbeziehung der Subdifferenziale und Normalenkegel zum Fréchet-Fall verwendet um entsprechende Aussagen zu erhalten.

Satz 9.3 (Clarke-Konzept) Die allgemeinen Voraussetzungen seien wie beim Fréchet-Konzept.

Vor.: Sei f außerdem lokal lipschitzstetig um x .

Beh.: $0 \in \partial^{Cl} f(x) + N_M^{Cl}(x)$

Bew.: Wir benutzen die Bedingung beim Fréchet-Konzept 2. Fall. Man braucht nur die Summenregel anzuwenden auf $f' = f + i_M$.

$$0 \in \partial^F f'(x) \subset \partial^{Cl} f'(x) \subset \partial^{Cl} f(x) + N_M^{Cl}(x).$$

Bem.: Diese Bedingung ist insbesondere relevant für konkrete Berechnungen mit PC^1 -Funktionen.

Satz 9.4 (Limiting-Fréchet-Konzept) Voraussetzungen seien wie beim Fréchet-Konzept.

Vor.: Sei f außerdem lokal lipschitzstetig um x .

Beh.: $0 \in \partial^{lim,F} f(x) + N_M^{lim,F}(x)$

Bew.: Konsequenz aus der Summenformel (6.1) und dem 2. Fall der Fréchet-Optimalitätsbedingungen angewandt auf die Funktion $f + i_M$, wobei i_M die Indikatorfunktion einer beliebigen abgeschlossenen Menge M ist.

9.1 „Gegenbeispiele“ für Optimalitätsbedingungen

Sehr interessant ist nun ein Vergleich der notwendigen Optimalitätsbedingungen einer differenzierbaren Funktion f auf den Mengen $M_4 = \{(x, y) | xy = 0\}$ oder $M_2 = \{(x, y) | xy = 0, (x, y) \geq 0\}$ im Punkt $(0, 0)$. Siehe auch Abschnitt 3.

Beispiel 16 (Optimalität auf M_4)

$$\begin{aligned} Cf() : & \pm Df(0, 0)e_i \geq 0 \Rightarrow Df(0, 0) = 0 \\ N_M^F : & Df(0, 0) = 0 \\ N_M^{lim,F} : & Df(0, 0)e_1 = 0 \text{ oder } Df(0, 0)e_2 = 0 \\ N_M^{Cl} : & Df(0, 0) \in \mathbb{R}^2 \text{ bzw. keine Bedingung} \end{aligned}$$

Beispiel 17 (Optimalität auf M_2)

$$\begin{aligned} Cf() : & Df(0, 0)v \geq 0 \forall v \in \mathbb{R}_+^2 \Rightarrow Df(0, 0) \in \mathbb{R}_+^2 \\ N_M^F : & -Df(0, 0) \in \mathbb{R}_-^2 \Rightarrow Df(0, 0) \in \mathbb{R}_+^2 \\ N_M^{lim,F} : & Df(0, 0) \in \mathbb{R}_+^2 \text{ oder } Df(0, 0)e_1 = 0 \text{ oder } Df(0, 0)e_2 = 0 \\ N_M^{Cl} : & Df(0, 0) \in \mathbb{R}^2 \text{ bzw. keine Bedingung} \end{aligned}$$

Kontingenz-Ableitung und Fréchet-Subdifferential liefern die selben Bedingungen aber die anderen Aussagen sind schwächer und können sogar unbrauchbar sein.

10 Konkrete Berechnungen

10.1 ...bei Differenzierbarkeit

Satz 10.1 (differenzierbarer Fall für $\partial^\alpha f(x)$) Ist f im Punkt x differenzierbar, so gilt

$$\partial^F f(x) = \{Df(x)\}.$$

Existiert die Ableitung in einer Umgebung und ist zusätzlich stetig in x , so folgt

$$\partial^{lim,F} f(x) = \partial^{Cl} f(x) = \{Df(x)\}.$$

Das heißt, die verallgemeinerten Subdifferentialiale sind gleich bzw. f ist regulär.

Beweis: $Df(x) \in \partial^F f(x)$ folgt aus der Definition. Weiterhin gilt

$$\begin{aligned} \partial^F f(x) &= \{v \mid \langle v, y - x \rangle \leq f(y) - f(x) + o(y - x)\} && y \in X \text{ beliebig, Gleichung 4.1} \\ &= \{v \mid \langle v, y - x \rangle - (f(y) - f(x)) \leq o(y - x)\} \\ &= \{v \mid \langle v - Df(x), y - x \rangle \leq o(y - x)\} && \text{Definition des Gradienten } Df \\ &\subset \{v \mid \langle v - Df(x), tu \rangle \leq o(tu)\} && t := \|y - x\| > 0, u := \frac{y-x}{t} \\ &= \left\{ v \mid \langle v - Df(x), u \rangle = \frac{o(tu)}{\|tu\|} \right\} && (-u) \text{ einsetzen, da } y \text{ bel.} \\ &= \{v \mid \langle v - Df(x), u \rangle = 0, u \in X^*, \|u\| = 1\} && t \rightarrow 0 \\ &= \{Df(x)\}. && \text{Einheitsvektoren einsetzen} \end{aligned} \tag{10.1}$$

Die Aussage für $\partial^{lim,F} f(x)$ ist eine direkte Konsequenz der Grenzwertbildung und der Stetigkeit der Ableitung. Die Aussage für $\partial^{Cl} f(x)$ folgt aus der Darstellung über Ableitungsgrenzwerte oder über konvexe Hülle von $\partial^{lim,F} f(x)$. \square

Bemerkung Die Umkehrung ($\partial^F f(x) = \{v\} \Rightarrow (f \text{ ist in } x \text{ differenzierbar.})$) gilt im Allgemeinen nicht, wie man im Beispielabschnitt 3 sieht an der Sägezahnfunktion zu $\frac{1}{2}|x|$. Auch die „einfachere“ Schnapperfunktion ist ein Gegenbeispiel.

Satz 10.2 (differenzierbarer Fall für Cf und Tf) Ist f um x lokal Lipschitzstetig und in x richtungsdifferenzierbar in Richtung u , so gilt $Cf(x; u) = \{f'(x; u)\}$. Ist f um x stetig differenzierbar, so gilt zusätzlich $Tf(x; u) = Df(x)u$.

Beweis:

$$\left\| \frac{f(x + t_k u_k) - f(x)}{t_k} - \frac{f(x + t_k u) - f(x)}{t_k} \right\| \leq L \|u_k - u\|,$$

wobei L eine lokale Lipschitzkonstante von f sei. Der Rest folgt direkt aus den Definitionen. \square

Satz 10.3 (differenzierbarer Fall für $N_{\text{Gph } g}^\alpha(x)$) Ist $g : \mathbb{R}^n \rightarrow \mathbb{R}^q$ im Punkt x differenzierbar, so zeigt man wie bei den Subdifferentialen

$$N_{\text{Gph } g}^F(x) = \left\{ (-Dg(x)^T v, v) \mid v \in \mathbb{R}^q \right\}.$$

Existiert die Ableitung in einer Umgebung und ist zusätzlich stetig in x , so gilt

$$N_{\text{Gph } g}^{lim,F}(x) = N_{\text{Gph } g}^{Cl}(x) = \left\{ (-Dg(x)^T v, v) \mid v \in \mathbb{R}^q \right\}.$$

Mithilfe der Definition des Co-Differentials folgt nun:

Folgerung 10.4 (differenzierbarer Fall für $D^\alpha g(x, g(x))y$) Ist $g : \mathbb{R}^n \rightarrow \mathbb{R}^q$ im Punkt x differenzierbar, so gilt

$$D^F g(x, g(x))y = \{Dg(x)^T y\}.$$

Existiert die Ableitung in einer Umgebung und ist zusätzlich stetig in x , so folgt

$$D^{lim, F} g(x, g(x))y = D^{Cl} g(x, g(x))y = \{Dg(x)^T y\}.$$

Satz 10.5 (differenzierbarer Fall für $N_M^\alpha(x)$) Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ in x differenzierbar, mit $Df(x) \neq 0$. Sei $M_1 = \{y | f(y) \leq 0\}$ und $M_2 = \{y | f(y) = 0\}$. Sofort aus den Definitionen folgt

$$N_{M_1}^F(x) \supset \mathbb{R}_+ Df(x) \text{ und } N_{M_2}^F(x) \supset \mathbb{R} Df(x).$$

Existiert die Ableitung in einer Umgebung und ist zusätzlich stetig in x , so haben wir wegen $Df \neq 0$ Regularität und es folgt mithilfe von Satz 8.1

$$N_{M_1}^\alpha(x) = \mathbb{R}_+ Df(x) \text{ und } N_{M_2}^\alpha(x) = \mathbb{R} Df(x) \quad \forall \alpha \in \{F; lim, F; Cl\}.$$

Bei mehreren Gleichungen brauchen wir die lineare Unabhängigkeit der Gradienten und die Elemente des Normalenkegel ergeben sich dann aus Linearkombinationen der Gradienten. Bei allen anderen Verallgemeinerungen braucht man härtere Bedingungen und M ist im Allgemeinen keine differenzierbare Mannigfaltigkeit mehr. Dieser Fall wird im folgenden Abschnitt nochmal wiederholt und war auch schon Bestandteil von Satz 8.1 und zugehöriger Folgerungen.

10.2 ...bei „Regularität“

10.2.1 Normalenkegel

Haben wir für eine Restriktionsmenge M eine Beschreibung der Form

$$M = \{x | g(x) \leq 0, h(x) = 0\},$$

wobei $h : \mathbb{R}^n \rightarrow \mathbb{R}^q$ und $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ stetig differenzierbare Funktionen seien, so wissen wir bereits:

$$N_M^F(x) \supset \left\{ Dh(x)^T u + Dg(x)^T v \mid u \in \mathbb{R}^q, v \in \mathbb{R}_+^p, v_i = 0 \quad \forall i : g_i(x) < 0 \right\}.$$

Haben wir nun ein u mit $Dh(x)u = 0$ und $g(x) + Dg(x)(u) \leq 0$, so ist es naheliegend zu fordern, dass u eine Richtung ist, in der es noch mehr zulässige Punkte $x_k \in M$ gibt. Diese Bedingung nennt sich auch Regularitätsbedingung und lautet konkret:

Definition 10.6 (Regularitätsbedingung) Eine Menge $M = \{x | g(x) \leq 0, h(x) = 0\}$, wobei $h : \mathbb{R}^n \rightarrow \mathbb{R}^q$, und $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ stetig differenzierbare Funktionen sind, heißt im Punkt $x \in M$ **regulär**, wenn aus $Dh(x)u = 0$ und $g(x) + Dg(x)(u) \leq 0$ ($u \neq 0$) folgt $\exists x_k \xrightarrow{x_k \in M} x, \frac{x_k - x}{\|x_k - x\|} \rightarrow \frac{u}{\|u\|}$, mit anderen Worten $u \in T_M^B(x)$.

Dies impliziert sofort

$$\{u \mid Dh(x)u = 0, Dg_i(x)u \leq 0 \forall i : g_i(x) = 0\} \subset T_M^B(x).$$

Über den Mittelwertsatz kann man auch die umgekehrte Teilmengenbeziehung zeigen. Durch Übergang zum Polarkegel erhält man dann den Normalenkegel.

Satz 10.7 ($N_M^F(x)$, wenn M regulär ist)

$$T_M^B(x) = \{u \mid Dh(x)u = 0, Dg_i(x)u \leq 0 \forall i : g_i(x) = 0\},$$

$$(N_M^B(x) =) N_M^F(x) = \left\{ Dh(x)^T u + Dg(x)^T v \mid u \in \mathbb{R}^q, v \in \mathbb{R}_+^p, v_i = 0 \forall i : g_i(x) < 0 \right\}.$$

Nun interessieren uns natürlich auch $N_M^{lim,F}(x)$ und $N_M^{Cl}(x)$. Dazu brauchen wir auch Regularität in einer Umgebung von x . Wir betrachten also u^k, v^k und $x_k \xrightarrow{x_k \in M} x$. Wir können davon ausgehen, dass x_k hinreichend nahe an x ist und somit $v_j^k = 0 \forall j : g_j(x) < 0$ gilt. Konvergiert nun u oder v , so ist klar, dass der Grenzwert in $N_M^F(x)$ liegt. Will man nun verhindern, dass u und v in der Norm gegen unendlich streben, damit keine schwer kalkulierbaren Grenzwerte dazukommen, so benötigt man stärkere Forderungen. Wie bereits gezeigt mit Satz 8.1, ist MFCQ eine solche Bedingung. Mit ihr erhält man dann normale Regularität und folglich, dass außer dem Kegel der konvexen Optimierung alle Normalenkegel zusammenfallen.

Ein Spezialfall entsteht, wenn man g weg lässt.

Folgerung 10.8 (differenzierbarer Fall für $N_{h^{-1}(0)}^\alpha(x)$) Ist $h : \mathbb{R}^n \rightarrow \mathbb{R}^q$ stetig differenzierbar um x und sind die Gradienten $Dh_i(x)$ linear unabhängig, so ist Regularität und sogar MFCQ erfüllt und es gilt

$$N_{h^{-1}(0)}^\alpha(x) = \left\{ (Dh(x)^T u) \mid u \in \mathbb{R}^q \right\} \forall \alpha \in \{F; lim, F; Cl\}.$$

10.2.2 Optimalität, Lagrange-Funktion, Karush-Kuhn-Tucker-Bedingung

Definition 10.9 (Standardaufgabe) Die Restriktionsfunktionen $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ und $h : \mathbb{R}^n \rightarrow \mathbb{R}^q$ und die Zielfunktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ seien stetig differenzierbar.

Die Aufgabe lautet dann

$$\begin{aligned} & \text{minimiere } f(x) \\ & \text{mit} \\ & g(x) \leq 0, \\ & h(x) = 0. \end{aligned}$$

Satz 10.10 (Optimalitätsbedingung) Sei x ein lokal minimaler Punkt. Die Restriktionsmenge $M = \{x \mid g(x) \leq 0, h(x) = 0\}$ sei regulär in $x \in M$ zum Beispiel mittels MFCQ. Mit Satz 9.2 Fall 1 und der obigen Formel für den Normalenkegel erhalten wir die Optimalitätsbedingung

$$\exists y \in \mathbb{R}_+^p, z \in \mathbb{R}^q : 0 = Df(x) + \sum_{i=1}^p Dg_i(x)y_i + \sum_{j=1}^q Dh_j(x)z_j \text{ und } y_i g_i(x) = 0 \forall i.$$

Diese Bedingung kann man nun unterschiedlich umschreiben. Dafür führt man eine „Hilfsfunktion“ ein, die für manche Beweise und allgemeine Betrachtungen wichtig ist.

Definition 10.11 (Lagrangefunktion)

$$L(x, y, z) = f(x) + \sum_{i=1, g_i(x)=0}^p g_i(x)y_i + \sum_{j=1}^q h_j(x)z_j$$

Die Optimalitätsbedingung ist dann $D_x L = 0$, auch Lagrangebedingung genannt. Die Zulässigkeit von x kann man mit $D_y L = 0$ und $D_z L = 0$ beschreiben. Die Bedingung $y_i g_i(x) = 0$ ist schon indirekt in L enthalten, da wir in der Summe nur die y_i betrachten mit $g_i(x) = 0$, was einem Nullsetzen der anderen y_i gleich kommt. Die Bedingung $y \in \mathbb{R}_+^p$ ist natürlich weiterhin zu fordern zum Beispiel über den Definitionsbereich. Die übliche andere Bedingung zur Lagrangebedingung ist jedoch die

Definition 10.12 (Komplementaritätsbedingung)

$$\begin{aligned} y &\geq 0, \\ -g(x) &\geq 0, \\ y_i g_i(x) &= 0 \quad \forall i. \end{aligned}$$

Teilweise taucht diese Bedingung zusätzlich mit $h(x) = 0$ auf.

Zusammen bilden die Lagrange- und die Komplementaritäts-Bedingung (mit der Bedingung $h(x) = 0$) die **Karush-Kuhn-Tucker-Bedingung**, welche meist auf KKT-Bedingung abgekürzt wird. Punkte, die diese Bedingung erfüllen, nennt man dann **KKT-Punkte**.

10.2.3 Die Ableitung der Kojima-Funktion

Die Kojimafunktion beschreibt die Lagrangegleichung und die Komplementaritätsbedingung durch ihre Nullstellen. Zu ihrer Beschreibung benötigen wir $y_i^+ := \max\{y_i, 0\}$ und $y_i^- := \min\{y_i, 0\}$.

Definition 10.13 (Kojima-Funktion) Seien $f : \mathbb{R}^n \rightarrow \mathbb{R}$ die Zielfunktion, $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ die Funktion für die Ungleichungsrestriktion $g(x) \leq 0$ und $h : \mathbb{R}^n \rightarrow \mathbb{R}^q$ für die Gleichungsrestriktion $h(x) = 0$. Alle 3 Funktionen seien lokal lipschitzstetig differenzierbar. Die zugehörige Kojimafunktion lautet dann

$$\begin{aligned} \Phi(x, y, z) &:= \begin{pmatrix} \Phi^1(x, y, z) \\ \Phi^2(x, y, z) \\ \Phi^3(x, y, z) \end{pmatrix} : \mathbb{R}^{n+p+q} \rightarrow \mathbb{R}^{n+p+q} \text{ mit} \\ \Phi^1(x, y, z) &:= (Df(x)^T + Dg(x)^T y^+ + Dh(x)^T z) \text{ und} \\ \Phi^2(x, y, z) &:= g(x) - y^- \\ \Phi^3(x, y, z) &:= h(x). \end{aligned}$$

Man sieht nun sofort

$$(x, y^+, z) \text{ ist ein KKT-Punkt} \Leftrightarrow \Phi(x, y^+ + g(x), z) = 0.$$

Außerdem erhält man mit

$$M(x) = \begin{pmatrix} Df & Dg_1 & \dots & Dg_p & 0_p^T & Dh_1 & \dots & Dh_q \\ g & 0_p & \dots & 0_p & -E_p & 0_p & \dots & 0_p \\ h & 0_q & \dots & 0_q & 0_{p*q} & 0_q & \dots & 0_q \end{pmatrix} \quad (10.2)$$

$$N(y, z) = (1, y_1^+, \dots, y_p^+, y_1^-, \dots, y_p^-, z_1, \dots, z_q)^T \quad (10.3)$$

$$\text{die Produktform} \quad (10.4)$$

$$\Phi(x, y, z) = M(x)N(y, z). \quad (10.5)$$

Damit die Matrix lesbar bleibt wurden die Gradienten wieder mit den Ableitungen identifiziert.

Jetzt kann man Satz 6.22 anwenden.

Folgerung 10.14 (Produktregel für die Kojimafunktion)

$$T\Phi((x, y, z); (u, v, w)) = TM(x; u)N(y, z) + M(x)TN((y, z); (v, w)), \quad (10.6)$$

$$C\Phi((x, y, z); (u, v, w)) = CM(x; u)N(y, z) + M(x)CN((y, z); (v, w)). \quad (10.7)$$

Die Ableitung von N kann man allerdings explizit ausrechnen.

Satz 10.15 (Ableitung von $N(y, z)$)

$$TN((y, z); (v, w)) = \left\{ \left(\begin{array}{c} 0 \\ (v_i r_i)_{i=1, \dots, p} \\ v_i(1 - r_i)_{i=1, \dots, p} \\ w \end{array} \right) \middle| r_i \in \begin{bmatrix} \{0\}, & y_i < 0 \\ [0; 1], & y_i = 0 \\ \{1\}, & y_i > 0 \end{bmatrix} \right\} \text{ und}$$

$$CN((y, z); (v, w)) = \left\{ \left(\begin{array}{c} 0 \\ (v_i r_i)_{i=1, \dots, p} \\ v_i(1 - r_i)_{i=1, \dots, p} \\ w \end{array} \right) \middle| r_i = \begin{bmatrix} 0, & y_i < 0 \\ 0, & y_i = 0, v_i \leq 0 \\ 1, & y_i = 0, v_i > 0 \\ 1, & y_i > 0 \end{bmatrix} \right\}$$

Satz 10.16 (C^2 -Fall) Sind nun f, g und h sogar zweifach stetig differenzierbar, so kann man die Ableitungen von M auch berechnen und erhält insgesamt, siehe [KK02] Kapitel 7.3,

$$C\Phi((x, y, z); (u, v, w)) = \{A_r(u, v, w)\} \text{ und } T\Phi((x, y, z); (u, v, w)) = \{A_r(u, v, w)\}.$$

Wir verwenden die zuvor beschriebenen möglichen r und definieren die dadurch verschiedenen

$$A_r = \begin{pmatrix} D_x^2 L(x, y^+, z) & r_1 Dg_1^T & \dots & r_p Dg_p^T & Dh^T \\ Dg_1 & -(1 - r_1) & & O & 0 \\ \vdots & & \ddots & & \vdots \\ Dg_m & O & & -(1 - r_p) & 0 \\ Dh & 0 & \dots & 0 & 0 \end{pmatrix},$$

wobei $L(x, y^+, z)$ die zu f, g, h gehörige Lagrangefunktion ist.

10.2.4 Die Ableitung der Funktion der stationären Punkte

Über die Kojima-Funktion bzw. die Definition des KKT-Punktes kann man stationäre Punkte definieren. Sei also $\Phi : \mathbb{R}^{n+p+q} \rightarrow \mathbb{R}^{n+p+q}$ wieder die zu stetig differenzierbaren f, g, h gehörende Kojima-Funktion.

Definition 10.17 (stationärer Punkt)

$$x \in \mathbb{R}^n \text{ ist ein stationärer Punkt} \Leftrightarrow \exists (y, z) \in \mathbb{R}^{p+q} : \Phi(x, y, z) = 0.$$

Nun interessieren uns Störungen dieses Systems in der rechten Seite.

Definition 10.18 (gestörte Aufgabe) Zu einer Störung $(a, b, c) \in \mathbb{R}^{n+p+q}$ betrachten wir die Aufgabe

$$\begin{aligned} &\text{minimiere } f(x) - \langle a, x \rangle \\ &\text{mit} \\ &g(x) \leq b, \\ &h(x) = c. \end{aligned}$$

Man sieht sofort, dass (x, y, z) eine Lösung zur zugehörigen Kojima-Funktion ist genau dann, wenn $\Phi(x, y, z) = p := (a, b, c)$. Wobei hier Φ die Kojimafunktion zur Standardaufgabe ist ohne Störung bzw. Störung Null.

Nun kann man die zur Störung gehörenden stationären Punkte untersuchen.

Definition 10.19 (Abbildung der stationären Punkte bei Störung der rechten Seite)

$$X(p) = \left\{ x \mid \exists (y, z) \in \mathbb{R}^{p+q} : \Phi(x, y, z) = p \right\}.$$

Dabei wird im Allgemeinen p nahe Null betrachtet. Damit die Menge der zu x gehörigen Lagrangemultiplikatoren (y, z) beschränkt bleibt und es überhaupt eine Verbindung zwischen $X(p)$ und optimalen Lösungen der Standardaufgabe bzw. der gestörten Aufgabe gibt, gehen wir hier davon aus, dass **MFCQ** erfüllt ist.

Per Definition entspricht $X(p)$ der Funktion $\Phi^{-1}(p)$ mit anschließender Projektion auf die x -Komponente. Die Untersuchung der Ableitung von $X(p)$ ist wie die Ableitung der Kojimafunktion wichtig, um Aussagen über das Verhalten der Lösungen machen zu können bei kleinen Störungen bzw. bei Änderung von Parametern. Die Untersuchung ist für $X(p)$ erschwert, da wir nur noch indirekt mit den Lagrangemultiplikatoren arbeiten können. Damit der Umgang mit den Ableitungen einfacher wird und wir schöne Formeln erhalten, fordere ich, dass f, g und h zweifach stetig differenzierbar sind. Mithilfe von Theorem 8.19 aus [KK02] erhalten wir dann folgenden Satz

Satz 10.20 (Kontingente Ableitung von $X(p)$) Seien alle zuvor genannten Bedingungen erfüllt. Sei $x \in X(p)$ und $\Phi(x, y, z) = M(x)N(y, z)$ die beschriebene Produktform der Kojima-Funktion. Dann gilt

$$\begin{aligned} &u \in CX((p, x); v) \Leftrightarrow \\ &v \in DM(x)uN(y', z') + M(x)CN((y', z'); \mathbb{R}^{p+q}) \text{ für ein } (y', z') : \Phi(x, y', z') = p. \end{aligned}$$

Die Formel von $CN((y', z'); (w_y, w_z))$ ist dafür im Abschnitt zur Ableitung der Kojima-Funktion zu finden.

Für die Thibault-Ableitung wurde bisher keine solche konkrete Formel gefunden, sondern nur eine Grenzwertbedingung, welche in Theorem 8.39 in [KK02] bewiesen wird.

Satz 10.21 (Thibault-Ableitung von $X(p)$) Zur Einfachheit fordern wir $g(x) = 0$ und betrachten die Aufgabe ohne h . Sei L die zugehörige Lagrangefunktion.

Dann gilt $u \in TX(p, x)(a, b)$ genau dann, wenn $\exists t_k \searrow 0, x_k \rightarrow x$ mit zugehörigen $y_k \rightarrow y, y'_k \rightarrow y'$ mit $\Phi(x, y) = \Phi(x, y') = p$, sodass

$$\begin{aligned} a &= \lim_{k \rightarrow \infty} D_{xx}^2 L(x, y') u + \frac{Dg(x_k)^T \alpha_k}{t_k} \\ b &= \lim_{k \rightarrow \infty} Dg(x) u - \frac{\beta_k}{t_k} \text{ wobei} \\ \alpha_k &= y'_k{}^+ - y_k{}^+ \\ \beta_k &= y'_k{}^- - y_k{}^- \end{aligned}$$

Im Allgemeinen gibt es keine konkrete Formel für die Berechnung der Thibault-Ableitung von $X(p)$, sondern nur in Spezialfällen, wie auch für das Co-Differential dieser Abbildung. Dies wird bewiesen in [KK05].

10.3 ...mit Maximumtypfunktionen

10.3.1 Definition und Subdifferential

Ist eine Funktion nicht mehr differenzierbar, so ist die Berechnung eines Subdifferentials etwas schwieriger. Um allerdings entsprechende Formeln angeben zu können, braucht man eine konkretere Beschreibung der Funktion. Als erste nehmen wir die Beschreibung für eine Maximumtypfunktion. Im nächsten Kapitel folgt eine etwas allgemeinere.

Definition 10.22 (Maximumtypfunktion)

$$F(x) = \max \{f_i(x) | i = 1, \dots, m\}, f_i : \mathbb{R}^n \rightarrow \mathbb{R}, \text{ stetig differenzierbar.}$$

Satz 10.23 (Subdifferential von Maximumtypfunktionen) Es gilt

$$\partial^F F(x) = \partial^{Cl} F(x) = \partial^{lim, F} F(x) = \text{conv} \{Df_i(x) | F(x) = f_i(x)\}.$$

Folglich sind Maximumtypfunktionen regulär wie auch deren Summen und positive Vielfache.

Beweisskizze: $Df_i(x) \in \partial^F F(x)$ ($F(x) = f_i(x)$) folgt sofort wegen

$$Df_i(x)(y - x) = f_i(y) - f_i(x) + o(y - x) \leq F(y) - F(x) + o(y - x).$$

Die andere Inklusion $\partial^{Cl} F(x) \subset \text{conv} \{Df_i(x) | F(x) = f_i(x)\}$ folgt aus der Darstellung von $\partial^{Cl} F(x)$ mittels Ableitungsgrenzwerten und weil alle Funktionen mit $f_j(x) < F(x)$ auch in einer Umgebung kleiner sind und somit weggelassen werden können bei der Betrachtung. Die Gleichung erhält man, weil $\partial^F F(x)$ konvex ist und durch die bekannten Teilmengenbeziehungen. \square

Ein Beispiel wäre hier die Betragsfunktion $|x| = \max\{x, -x\}$.

Im folgenden Abschnitt zu PC^1 -Funktionen werden wir zeigen, dass die Maximumtypfunktionen lokal Lipschitzstetig sind. Somit können wir Satz 5.11 und die Subdifferentialformel anwenden.

Folgerung 10.24 (Maximumtypfunktion & Normalenkegel) Sei $M = \{x | F(x) \leq 0\}$ und gelte $0 \notin \text{conv}\{Df_i(x) | F(x) = f_i(x)\}$, was nichts anderes bedeutet, als dass die aktiven Ableitungen positiv linear unabhängig sind. Dann gilt

$$N_M^F(x) = N_M^{\text{lim},F}(x) = N_M^{Cl}(x) = \mathbb{R}_+ \text{conv}\{Df_i(x) | F(x) = f_i(x)\}.$$

Schaut man genauer auf diesen Satz, so ist er nichts anderes als Anwendung der MFCQ auf die Menge $M = \{x | g_i(x) \leq 0 \forall i\}$.

Bemerkung Minimumtypfunktion

Die Minimumtypfunktion ist das negative der Maximumtypfunktion. Somit ist die Folgerung auch gültig für Minimumfunktionen, wenn man dann $M = \{x | F(x) \geq 0\}$ betrachtet, wobei man dann jedoch die negativen Ableitungen verwendet.

Im Subdifferential war die Maximeigenschaft essentiell wichtig, wie man zum Beispiel an der Funktion $-|x|$ sieht. Minimumtypfunktionen sind ein Spezialfall der im nächsten Kapitel untersuchten PC^1 -Funktionen.

10.3.2 Die Kontingent-Ableitung des Subdifferentials

Im Abschnitt zu PC^1 -Funktionen werden wir auf den Effekt treffen, dass die Richtungsableitung wieder eine PC^1 -Funktion ist und wieder richtungsdifferenzierbar ist.

In diesem Abschnitt werden wir uns mit einer anderen Art der zweiten Ableitung beschäftigen, für die es noch eine Formel gibt. Wir beschränken uns dabei auf den Fall einer Maximumtypfunktion, $F(x) = \max\{f_i(x) | i = 1, \dots, m\}$, wobei die $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ zweifach stetig differenzierbar sind. Es gibt zwar auch Gesetzmäßigkeiten für allgemeinere Funktionen, die man aus [RW98] herleiten kann. Dazu braucht man dort Theorem 13.14 und 13.40. Für diese Funktionen werden aber keine konkreten Formeln berechnet im Gegensatz zu unserem Fall, wo in [KK02] Theorem 9.7 folgender Satz zu finden ist.

Satz 10.25 ($C(\partial^\alpha F)((x, x^*); u)$) Es gilt $v \in C(\partial^\alpha F)((x, x^*); u)$ genau dann, wenn $\exists \alpha, y \in \mathbb{R}_+^m$ und $z \in \mathbb{R}^n$, sodass $(0) \emptyset \neq J := \{i | \alpha_i + y_i > 0\} \subset I_F(x) := \{i | F(x) = f_i(x)\}$,

- (1) $\langle Df_i(x) - x^*, u \rangle \leq 0 \forall i \in I_F(x)$,
- (2) $\langle Df_i(x) - x^*, u \rangle = 0 \forall i \in J$,
- (3) $\langle Df_i(x) - x^*, z \rangle + \langle u, D^2 f_i(x)u \rangle \leq \langle v, u \rangle$ falls $\langle Df_i(x) - x^*, u \rangle = 0$ und $i \in I_F(x)$ und Gleichheit gilt für $i \in J$,
- (4) $x^* = \sum_{i \in J} \alpha_i Df_i(x)$, $\sum_{i \in J} \alpha_i = 1$,
- (5) $v = \sum_{i \in J} \alpha_i D^2 f_i(x)u + \sum_{i \in J} y_i (Df_i(x) - x^*)$,

Hier werden die Df_i mit den Gradienten identifiziert. $I_F(x)$ ist die Menge der aktiven Indizes von F im Punkt x .

Bemerkung (zum Satz) Der Beweis in [KK02] reduziert das Problem zunächst, indem es x^* zu Null macht und zusätzlich einen quadratischen Term eliminiert. Im Beweis werden dort unter anderem Ekeland's Variationsprinzip angewendet und lineare Dualität.

Nun werde ich auf die einzelnen Größen und Bedingungen noch etwas näher eingehen.

J bezeichnet so etwas wie die aktiven Indizes, die jeweils betrachtet werden, welches nicht unbedingt alle aktiven Indizes sein müssen.

α beschreibt x^* als Konvexkombination aktiver Gradienten und Bedingung (4) sagt dann nichts anderes als $x^* \in \partial^\alpha F(x) = \text{conv} \{Df_i(x) | F(x) = f_i(x)\}$.

Bedingung (1) stellt eine spezielle Forderung an x^* und u . Ist diese nicht erfüllt, so ist die Kontingent-Ableitung leer. Umgeschrieben bedeutet die Bedingung, dass u keine Anstiegsrichtung von $G(x') = F(x') - \langle x^*, x' - x \rangle$ ist.

Aufbauend darauf sagt Bedingung (2), dass für die Indizes, die in der Linearkombination von x^* genommen werden, die Ableitung der Funktionen G in Richtung u Null zu sein hat. Außerdem werden alle Indizes, bei denen dies zutrifft, potentiell mit einem positiven Faktor y_i in die Konstruktion von v einbezogen.

In welcher Form dies geschieht, legt dann Bedingung (4) fest. Hier hat man eine Linearkombination der Ableitungen von G in Richtung u und der Richtungsableitung von $H(x') := \sum_{i \in J} \alpha_i Df_i(x')$, einer Art linearen Erweiterung von x^* .

Die einzig große Forderung ist also Bedingung (3). Diese ist allerdings nichts anderes als lineare Dualität angewandt auf die lineare Aufgabe $(LP)_u$ mit $K := \{i \in I_F(x) | \langle Df_i, \cdot \rangle\}$:

$$\max_{y'} \left\{ \sum_{i \in K} \langle u, D^2 f_i(x) u \rangle y_i \mid \sum_{i \in K} (Df_i(x) - x^*) y_i = 0, \sum_{i \in K} y_i = 1, y_i \geq 0 \forall i \in K - J \right\}.$$

Hierbei ist $\alpha = y'$ die Lösung, $\langle v, u \rangle$ der optimale Wert und z eine entsprechende duale Variable.

Diese Aufgabe scheint in [RW98] der Formel zu Beispiel 13.16, Seite 600 zu entsprechen. v ist dort unser x^* und w unser u . Der interessante Punkt dabei ist, dass mit Theorem 13.40 ein Zusammenhang hergestellt wird zwischen unserer berechneten Ableitung und dem Subdifferential der Optimalwertfunktion zur Aufgabe.

Offen bleibt die Frage, ob diese Zusammenhänge womöglich neue Erkenntnisse bringen.

10.4 ...mit stückweise differenzierbaren Funktionen

10.4.1 Definitionen und Stetigkeit

Für eine etwas allgemeinere Beschreibung als die für die Maximumtypfunktion brauchen wir den Begriff der stückweise differenzierbaren Funktion.

Definition 10.26 (stückweise differenzierbar) Sei $F : \mathbb{R}^n \rightarrow \mathbb{R}$ eine stetige Funktion. Wir nennen F stückweise differenzierbar oder auch PC^1 -Funktion, wenn folgendes gilt:

Für jedes x existiert eine Umgebung U und endlich viele stetig differenzierbare Funktionen $f_i : U \rightarrow \mathbb{R}$ $i = 1, \dots, m$, sodass gilt

$$F(z) \in \{f_1(z), \dots, f_m(z)\} \forall z \in U.$$

Analog definieren wir für Abbildungen $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ mit größerem Wertebereich.

Beispiele für Funktionen, für die solch eine Beschreibung existiert, wären Maximumtypfunktionen, Minimumtypfunktionen, wie die negative Betragsfunktion, und Beispiel 7. Die Sägezahnfunktionen sind keine PC^1 -Funktionen, da die aktiven f_i immer wieder den Anstieg (1 bzw. -1) und 0 haben würden und zwar um so öfter, je näher man an den Nullpunkt kommt. Folglich kann die Ableitung der f_i nicht stetig sein, da die Ableitung im Nullpunkt nicht gleichzeitig 1 und 0 sein kann.

Als nötige Hilfsfunktion benötigen wir den Begriff der wesentlich aktiven Indizes.

Definition 10.27 (aktive Indizes bei PC^1 -Funktionen) $I_F(x) := \{i | F(x) = f_i(x)\}$ ist die Menge der aktiven Indizes. Die Menge der wesentlich aktiven Indizes ist

$$I_F^w(x) := \left\{ i \mid x \in \overline{\text{int} \{z \in U \mid F(z) = f_i(z)\}} \right\},$$

welche etwas kleiner ist, aber nie leer. Entsprechend ihrer Indizes sind dann auch die Funktionen f_i aktiv bzw. wesentlich aktiv in x .

Bemerkung Da die Ermittlung der wesentlich aktiven Indizes sehr aufwendig sein kann, sollte man, wenn möglich, darauf achten, dass die Funktionen f_i so gewählt werden, dass $I_F^w(x) = I_F(x)$ für jedes x gilt. Dies ist lokal immer möglich, wie in [Sch94] Proposition 4.1.1 bewiesen. Der Grund ist, dass ein Index i , der in x nicht wesentlich aktiv ist, auch in einer Umgebung O von x nicht wesentlich aktiv ist und somit kann man $O \cap U$ statt der Umgebung U in der Definition von PC^1 -Funktionen betrachten und f_i streichen.

Beispiel 18 (unterschiedliche Indexfunktionen) Sei $F(x) = \max \{x^2, x^4\}$. Dann gilt

$$I_F(0) = \{1, 2\} \text{ bzw. } \{i_{x^2}, i_{x^4}\}, I_F(1) = \{1, 2\}, I_F^w(0) = \{1\} \text{ und } I_F^w(1) = \{1, 2\}.$$

Selbst bei Maximumtypfunktionen können sich die beiden Indexfunktionen unterscheiden, wie man an diesem Beispiel sieht.

Lemma 10.28 (Stetigkeit der Indexfunktion) Die Abbildung $I_F^w(\cdot) : U \rightrightarrows \{1, \dots, m\}$ ist abgeschlossen und semiunterhalbstetig, was hier heißen soll

$$\forall i \in I_F^w(x) \exists x_k \rightarrow x, i \in I_F^w(x_k) \forall k.$$

Beweis: Semiunterhalbstetigkeit folgt direkt aus der Definition der wesentlich aktiven Indizes, indem wir $x_k \in \text{int} \{z \in U \mid F(z) = f_i(z)\}$ wählen.

Für den Beweis der Abgeschlossenheit gelte wieder oBdA $I_F(x) = I_F^w(x)$. Nun betrachten wir eine Folge $x_k \rightarrow x$ und oBdA gelte $i \in I_F^w(x_k) \forall k$. Nun gilt wegen der Stetigkeit von F und f_i : $F(x) = \lim F(x_k) = \lim f_i(x_k) = f_i(x)$ und somit $i \in I_F^w(x)$. \square

Lemma 10.29 (Lipschitzstetigkeit der PC^1 -Funktion) F ist um x lokal lipschitzstetig.

Beweis: Sei $O \subset U$ eine kompakte konvexe Umgebung von x . Da die Norm-Abbildung und die Ableitungen der f_i stetig sind, hat $\|Df_i(y)\|$ auf O ein Maximum L_i . Sei L das Maximum aller L_i . Wir rechnen aus für $y, z \in O$

$$\|f_i(y) - f_i(z)\| = \left\| \int_0^1 Df_i(y + t(z - y))(z - y) dt \right\| \leq L \|y - z\|.$$

OBdA seien alle aktiven Indizes wieder wesentlich. Aufgrund der Stetigkeit der Funktion der wesentlich aktiven Indizes können wir auf der Strecke von y nach z endlich viele a_j auswählen, sodass gilt $\exists i_j \in I_F^w(a_j) \cap I_F^w(a_{j+1})$, wobei $a_0 := y$ und $a_s := z$. Natürlich sollen die a_j „aufeinanderfolgend“ auf der Strecke liegen, wodurch in der Dreiecksungleichung die bekannte Gleichheit gilt. Nun haben wir:

$$\|F(y) - F(z)\| \leq \sum_{l=0}^s \|f_{i_l}(a_l) - f_{i_l}(a_{l+1})\| \leq L \sum_{l=0}^s \|a_l - a_{l+1}\| = L \|y - z\|,$$

Somit ist L die gesuchte Lipschitzkonstante auf der Umgebung um x . \square

Man kann auch genauso zeigen, dass jede stetige Auswahl von Lipschitzfunktionen wieder lipschitzstetig ist mit einer Lipschitzkonstante die dem Maximum der einzelnen Lipschitzkonstanten entspricht. Die beiden vorangegangenen Sätze sind unter anderem in [Sch94] zu finden, wo die Beweise etwas ausführlicher sind.

10.4.2 Subdifferential und Richtungsableitung

In [Kum88] bzw. [Sch94] gibt es folgende nützliche Formel, deren Beweis wir mit den vorangegangenen Hilfssätzen führen.

Satz 10.30 (Subdifferential von PC^1 -Funktionen) Es gilt:

$$\{Df_i(x) | i \in I_F^w(x)\} \cup \partial^F F(x) \subset \partial^{lim, F} F(x) \text{ und } \text{conv} \{Df_i(x) | i \in I_F^w(x)\} = \partial^{Cl} f(x).$$

Beweis: Wir wissen bereits aus Satz 10.1

$$\partial^F F(z) = \partial^{lim, F} F(z) = \partial^{Cl} F(z) = \{Df_i(z)\} \forall z \in \text{int} \{z \in U | F(z) = f_i(z)\}.$$

Somit haben wir sofort durch Grenzwertbildung

$$\{Df_i(x) | i \in I_F^w(x)\} \cup \partial^F F(x) \subset \partial^{lim, F} F(x) \text{ und } \text{conv} \{Df_i(x) | i \in I_F^w(x)\} \subset \partial^{Cl} F(x),$$

da das Clarke-Subdifferential das Limiting-Fréchet-Subdifferential enthält und konvex ist. Mit Gleichung 4.4 und der lokalen Lipschitzstetigkeit von F um x genügt zu zeigen

$$\left\{ v = \lim_{k \rightarrow \infty} DF(x_k) \mid x_k \rightarrow x, DF(x_k) \text{ existiert} \right\} \subset \{Df_i(x) | i \in I_F^w(x)\}.$$

Sei also $v = \lim DF(x_k)$ und $x_k \rightarrow x$ die zugehörige Punktfolge. Sei $u \in \mathbb{R}^n$ eine beliebige Richtung, in die wir die Ableitung von F im Punkt x_k betrachten wollen.

Wie bereits zuvor erwähnt, seien oBdA alle aktiven Indizes wesentlich. Aufgrund der Abgeschlossenheit der Indexfunktion gilt oBdA $I_F^w(x_k + tu) \subset I_F^w(x_k) \subset I_F^w(x)$ für hinreichend kleine t . Da die Menge der aktiven Indizes endlich ist, existiert nun eine Folge $t_l \searrow 0$ und ein $i \in I_F^w(x_k)$, sodass $F(x_k + t_l u) = f_i(x_k + t_l u)$. Die Folge erhält man, indem man zunächst eine beliebige Folge nimmt von aktiven Indizes zu beliebiger Folge t_l bzw. $x_k + t_l u$ und dann eine Teilfolge betrachtet mit konstantem $i \in I_F^w(x_k + t_l u)$. Aufgrund der Stetigkeit ist i auch in x_k wesentlich aktiv. Nun gilt

$$DF(x_k)u = \lim_{t_l \searrow 0} \frac{F(x_k + t_l u) - F(x_k)}{t_l} = \lim_{t_l \searrow 0} \frac{f_i(x_k + t_l u) - f_i(x_k)}{t_l} = Df_i(x_k)u.$$

Wegen der Linearität des Differentials folgt, dass $V_i = \{v \mid DF(x_k)v = Df_i(x_k)v\}$ ein linearer Untervektorraum ist. Zusammen mit der zuvor erhaltenen Gleichung und da u dort beliebig war, erhalten wir die Gleichung

$$\mathbb{R}^n = \bigcup_{i \in I_F^w(x_k)} V_i \quad .$$

Da nun aber die V_i Untervektorräume sind, folgt nach der linearen Algebra, dass ein $i \in I_F^w(x_k)$ existiert mit $V_i = \mathbb{R}^n$ und somit $DF(x_k) = Df_i(x_k)$.

Betrachten wir nun die konvergente Folge $DF(x_k) = Df_{i(x_k)}(x_k)$, so kann man wieder zu einer konvergenten Teilfolge übergehen mit konstantem $i(x_k) = i \in I_F^w(x)$. Nun gilt

$$v = \lim DF(x_k) = \lim Df_i(x_k) = Df_i(x),$$

da die f_i stetig differenzierbar sind. Somit gilt $v \in \{Df_i(x) \mid i \in I_F^w(x)\}$. □

Bemerkung Der Beweis zeigt, dass noch eine etwas schärfere Bedingung gilt ohne konvexe Hülle:

$$\{v = \lim DF(x_k) \mid x_k \rightarrow x, DF(x_k) \text{ existiert}\} = \{Df_i(x) \mid i \in I_F^w(x)\}.$$

Die Vermutung

$$\{Df_i(x) \mid i \in I_F^w(x)\} \cup \partial^F f(x) = \partial^{lim, F} f(x)$$

ist im Allgemeinen falsch, wie man am recht einfachen Beispiel 7 sieht. Hier ist das Fréchet-Subdifferential einelementig, enthält aber keine der Ableitungen der aktiven Funktionen. Im Limiting-Fall erhalten wir die Vereinigung von zwei Mengen, die aus der konvexen Hülle von zwei unterschiedlichen Paaren von Ableitungen der aktiven Funktionen kommen. Dies zeigt, dass konkrete Formeln für die Berechnung der beiden Subdifferenziale nicht zu erwarten sind.

Für die Berechnung des Subdifferentials waren zwar nur die wesentlich aktiven Indizes interessant, aber für die Berechnung bei Maximumtypfunktionen kann man, wie bereits bewiesen, auch alle aktiven Indizes verwenden.

Folgerung 10.31 Analog für $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ zeigt man

$$\text{conv} \{Df_i(x) | i \in I_F^w(x)\} = \partial^{Cl} f(x),$$

wobei hier die verallgemeinerte Clarke-Ableitung gemeint ist.

Wir haben gerade im Beweis für die Formel des Subdifferential festgestellt, dass eine PC^1 -Funktion in differenzierbaren Punkten dem Differential einer ihrer wesentlich aktiven Funktionen entspricht. Dazu hatten wir die Richtungsableitung betrachtet und festgestellt, dass sie der Richtungsableitung einer wesentlich aktiven Funktion entspricht. Nun wäre natürlich von Interesse, wie die Richtungsableitung allgemein aussieht.

Satz 10.32 (Richtungsableitung einer PC^1 -Funktion F)

$F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ist richtungsdifferenzierbar und es gilt

$$F'(x; u) = Df_i(x)u \in \{Df_i(x)u | i \in I_F^w(x), \exists t_k \searrow 0, i \in I_F^w(x + tu)\}.$$

Aus der lokalen Lipschitzstetigkeit von F folgt sofort wegen der Existenz der Grenzwerte der Differenzenquotienten, dass $F'(x, \cdot)$ in der zweiten Komponente lipschitzstetig ist.

Beweis: Im folgenden werden sehr aktiv die Stetigkeitseigenschaften der Funktion der wesentlich aktiven Indizes benutzt. Desweiteren seien alle betrachteten aktiven Indizes wesentlich aktive Indizes.

u sei die Richtung in die wir ableiten wollen.

Für die Richtungsableitung müssen wir $F(x + t_k u)$ betrachten mit Variable t , welches auch eine PC^1 -Funktion ist. Folglich sei oBdA $u = 1$, $x = 0$ und $F = F(t)$, d.h. $n = 1$. Nun gilt oBdA :

$$\exists t_k^1 \searrow 0, 1 \in I_F^w(t_k^1) \text{ und somit } Df_1(0) = \lim_{k \rightarrow \infty} \frac{f_1(t_k^1) - f_1(0)}{t_k^1} \in CF(0; 1).$$

Jede andere Folge $s_k^1 \searrow 0$ würde aufgrund der Differenzierbarkeit der f_1 den selben Grenzwert liefern. Nun haben wir das erste Element der zu betrachtenden Menge

$$\left\{ v = \lim_{k \rightarrow \infty} \frac{F(t_k) - F(0)}{t_k} \mid t_k \searrow 0 \right\}$$

und müssen nach weiteren schauen bzw. zeigen, dass es keine weiteren gibt.

Ist nur der Index 1 für positive t wesentlich aktiv, wenn man sich 0 nähert, so sind wir fertig. Angenommen also, es existieren andere wesentliche aktive Indizes. Aufgrund der „Stetigkeit“ der Indexfunktion und weil sie nur endlich viele Werte annehmen kann, gilt wieder oBdA:

$$\exists t_k^2 \searrow 0, \{1, 2\} \subset I_F^w(t_k^2) \text{ und somit } Df_1(0) = Df_2(0).$$

Wieder wären wir fertig, wenn keine anderen Indizes wesentlich aktiv werden, wenn wir uns mit positiven t der 0 nähern. Sei dies wiederum nicht der Fall, dann folgt mit der selbigen Begründung:

$$\exists t_k^3 \searrow 0, \{1, 3\} \subset I_F^w(t_k^3) \vee \{2, 3\} \subset I_F^w(t_k^1) \text{ und somit } Df_1(0) = Df_2(0) = Df_3(0).$$

Analog zu der letzten Argumentation kann man nun fortfahren und zeigt in endlich vielen Schritten, dass alle Folgen $t_k \searrow 0$ mit $\exists i \in I_F^w(t_k) \forall k$ den selben Grenzwert für die Richtungsableitung liefern, da wir in unserem Algorithmus alle diese Folgen betrachtet haben und gezeigt haben, dass sie den selben Grenzwert liefern.

Der Trick besteht im wesentlichen darin, dass ein Index an den Grenzen der abgeschlossenen Intervalle, wo er wesentlich aktiv ist in \mathbb{R} , mit anderen Indizes zusammenfallen muss. Ausgehend von einem Index haben wir eine Kette von Indizes gebildet, die alle über unendlich viele Intervalle miteinander verbunden sind, wodurch wir letztendlich alle für die Richtungsableitung in 0 interessanten Indizes betrachten. □

Bemerkung In [Sch94] wird ähnlich verfahren, aber über den Hauptsatz der Integralrechnung und die Stetigkeit der Ableitung der Df_i argumentiert, während ich einen induktiven Beweis über die Indizes geführt habe, der die „Stetigkeit“ der Indexfunktion benutzt.

10.4.3 Weitere Ableitungen

Folgerung 10.33 (Kontingenz-Ableitung der PC^1 -Funktion F) Da $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ lokal lipschitzstetig ist gilt aufgrund des Satzes zur Richtungsableitung von PC^1 -Funktionen:

$$CF(x; u) = \{F'(x, u)\}.$$

Folgerung 10.34 (Fréchet-Subdifferential) Mit Satz 5.14 folgt für $F : \mathbb{R}^n \rightarrow \mathbb{R}$

$$(\partial^F F(\cdot))(x) = \partial^{ko} F'(x; \cdot)(0).$$

Satz 10.35 (Thibault-Ableitung der PC^1 -Funktion F) Für $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ haben wir mit Satz 5.19

$$\text{conv}(TF(x; u)) = \text{conv}(\{Df_i(x)u \mid i \in I_F^w(x)\})$$

und für $F : \mathbb{R}^n \rightarrow \mathbb{R}$ sogar

$$TF(x; u) = \text{conv}(\{Df_i(x)u \mid i \in I_F^w(x)\}).$$

Bemerkung Aufgrund der Richtungs-differenzierbarkeit sieht man nun leicht die Formel

$$\text{conv}(TF(x; u)) = \text{conv} \left(\limsup_{x_k \rightarrow x} Cf(x_k; u) \right),$$

wobei wir im eindimensionalen Fall wieder bei der Thibault-Ableitung die konvexe Hülle weglassen können.

Vergleicht man den $\lim \sup$ mit dem Limiting-Fréchet-Konzept, so wäre hier die Thibault-Ableitung das Analogon zum Clarke-Subdifferential, da die Kontingenz-Ableitung dem Fréchet-Konzept zuzuordnen ist und sich die Thibault-Ableitung aus $\lim \sup$ und konvexer Hülle ergibt, wie auch das Clarke-Subdifferential aus dem Fréchet-Subdifferential.

Aus dem Lemma 5.21 und den Sätzen 10.30 und 10.4 erhalten wir nun eine Formel für das Co-Differential.

Folgerung 10.36 (Co-Differential von PC^1 -Funktionen) Sei $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine um x lokal lipschitzstetige Funktion. Dann gilt

$$\text{conv}(D^{\text{lim},F}F(x, F(x))(v)) = \left\{ A^T v \mid A \in \partial^{Cl}(F)(x) \right\},$$

wobei hier die verallgemeinerte Jakobimatrix verwendet wird zur einfacheren Notation. Ist F zusätzlich eine PC^1 -Funktion, so haben wir

$$\left\{ Df_i(x)^T v \mid i \in I_F^w(x) \right\} \cup D^F F(x, F(x))(v) \subset D^{\text{lim},F}F(x, F(x))(v)$$

$$\text{und } \text{conv} \left\{ Df_i(x)^T v \mid i \in I_F^w(x) \right\} = \text{conv}(D^{\text{lim},F}F(x, F(x))(v)) \subset D^{Cl}F(x, F(x))(v).$$

Beweis: Die erste Gleichung folgt durch Anwendung der konvexen Hülle auf Lemma 5.21 und anschließendes anwenden von Satz 5.5 und Gleichung 4.4.

Die Zweite Gleichung wird wie im vorletzten Satz durch Grenzwertbildung bewiesen, allerdings mit der Gleichung $D^F g(x, g(x))y = \left\{ Dg(x)^T y \right\}$ für Differenzierbarkeitspunkte aus Folgerung 10.4.

Für die dritte Gleichung müssen wir wie zuvor zeigen, dass in Differenzierbarkeitspunkten von F die Ableitung mit einer der wesentlich aktiven f_i zusammenfällt und setzen dies in die erste Gleichung ein. Der letzte Teil folgt über die Gleichung (5.1): $N_M^{Cl}(x') = \text{conv}(N_M^{\text{lim},F}(x'))$. \square

10.4.4 Anwendung auf Normalenkegel und Optimalität

Nun werden wir Satz 5.11 (Beziehung Subdifferential und Clarke-Normalenkegel) und 10.5 (Normalenkegel bei stetig differenzierbarer Umgebung) auf die Formel für das Subdifferential aus Satz 10.30 anwenden.

Satz 10.37 (Normalenkegel zu PC^1 -Funktion) Sei $F : \mathbb{R}^n \rightarrow \mathbb{R}$ eine PC^1 -Funktion, die um x beschrieben ist wie in der Definition. Die zu betrachtenden Mengen sind $M_1 = \{y \mid F(y) \leq 0\}$ und $M_2 = \{y \mid F(y) = 0\}$ und es gelte $F(x) = 0$. Dann haben wir

$$\mathbb{R}_+ \left\{ Df_i(x) \mid i \in I_F^w(x) \right\} \cup N_{M_1}^F(x) \subset N_{M_1}^{\text{lim},F}(x) \text{ und}$$

$$\mathbb{R} \left\{ Df_i(x) \mid i \in I_F^w(x) \right\} \cup N_{M_2}^F(x) \subset N_{M_2}^{\text{lim},F}(x).$$

Gilt weiterhin $0 \notin \text{conv} \left\{ Df_i(x) \mid i \in I_F^w(x) \right\}$ bzw. dass die unterschiedlichen Ableitungen der wesentlich aktiven Funktionen positiv linear unabhängig sind, dann gilt

$$\begin{aligned} \text{conv}(\mathbb{R}_+ \left\{ Df_i(x) \mid i \in I_F^w(x) \right\}) &= N_{M_1}^{Cl}(x) \text{ und} \\ \text{conv}(\mathbb{R} \left\{ Df_i(x) \mid i \in I_F^w(x) \right\}) &= N_{M_2}^{Cl}(x). \end{aligned}$$

Beweis: Die ersten zwei Teilmengenbeziehungen sind eine direkte Konsequenz des Grenzwertübergangs in den Gleichungen aus Satz 10.5 .

Die darauf folgenden „ \subset “-Beziehungen folgen direkt aus $N_M^{Cl}(x) = \text{conv } N_M^{lim,F}(x)$ (Gleichung (5.1)) und den beiden anderen Teilmengenbeziehungen.

Die erste „ \supset “-Beziehung gilt wegen mittels Satz 5.11 und der Formel für das Clarke-Subdifferential von PC^1 -Funktionen.

Die zweite „ \supset “-Beziehung ist etwas aufwendiger und wird im folgenden bewiesen.

Wir definieren zunächst $C_1 = \mathbb{R}^n \times \{0\}$ und $C_2 = \text{gph } F$. Somit gilt

$$M_2 \times 0 = C_1 \cap C_2.$$

Wir erhalten somit aus der Rechenregel bei Mengenschnitt 8.4

$$N_{M_2 \times \{0\}}^{lim,F}(x, 0) \subset \{0\}^n \times \mathbb{R} + N_{\text{gph } f}^{lim,F}(x, 0),$$

unter der Bedingung $(0, t) \in N_{\text{gph } F}^{lim,F}(x, 0) \Rightarrow t = 0$.

Zunächst zeigen wir, dass diese Bedingung erfüllt ist und werden mit der Formel weiter rechnen. Es genügt die Fälle $t = 1$ und $t = -1$ auszuschließen, da wir mit Kegeln arbeiten. Die Bedingung lautet nun

$$0 \notin D^{lim,F} F(x, 0)(\pm 1).$$

Mit der vorigen Formel zum Co-Differential und unseren Voraussetzungen wissen wir jedoch

$$\text{conv}(D^{lim,F} F(x, 0)(\pm 1)) = \pm \text{conv} \{Df_i(x) | i \in I_F^w(x)\} \neq 0.$$

Folglich ist die Bedingung erfüllt.

Nun rechnen wir weiter. Dafür verwenden wir die Kreuzproduktregel 6.1:

$$N_{M_2 \times \{0\}}^{lim,F}(x, 0) = N_{M_2}^{lim,F}(x) \times \mathbb{R}.$$

Dies setzen wir in die erste Formel ein, deren Bedingung ja erfüllt ist und verwenden die Definition des Co-Differentials:

$$N_{M_2}^{lim,F}(x) \subset D^{lim,F} F(x, 0)(\mathbb{R}).$$

Die Skalierbarkeit des Co-Differentials und der Übergang zur konvexen Hülle mit Gleichung (5.1) liefert

$$N_{M_2}^{Cl}(x) \subset \text{conv}(\mathbb{R}_+ D^{lim,F} F(x, 0)(1)) + \text{conv}(\mathbb{R}_+ D^{lim,F} F(x, 0)(-1)) + D^{lim,F} F(x, 0)(0).$$

Und mit der eben benutzten Formel für das Co-Differential erhalten wir, weil die betrachteten Mengen alle konvexe Kegel sind:

$$N_{M_2}^{Cl}(x) \subset \text{conv}(\mathbb{R}_+ \{Df_i(x) | i \in I_F^w(x)\} + \mathbb{R}_- \{Df_i(x) | i \in I_F^w(x)\}) + 0$$

und somit die letzte zu zeigende Teilmengenbeziehung. \square

Anwendungsbeispiele wären $M_1 = \text{epi}(-|\cdot|)$ und $M_2 = \text{gph}(-|\cdot|)$ für diesen Satz.

Bemerkung ($0 \notin \text{conv} \{Df_i(x) | i \in I_F^w(x)\}$) Da wir eine konvexe Menge betrachten und sogar alle potentiellen Ecken kennen, können wir mit Hilfe des Trennungssatzes die Bedingung umformulieren. Wir erhalten

$$\exists u_0 : Df_i(x)(u_0) < 0 \text{ bzw. } \langle Df_i(x), u_0 \rangle < 0 \quad \forall i \in I_F^w(x).$$

Dies ähnelt der MFCQ bei Ungleichungsrestriktionen, allerdings resultiert die Bedingung vielmehr daher, dass bei nur einer Restriktion die Ableitung der Restriktionsfunktion nicht Null sein darf, egal ob wir eine Gleichung oder Ungleichung betrachten. Bei uns ist allerdings die Ableitung durch das Clarke-Subdifferential ersetzt.

Bemerkung (kompliziertere Mengen) Wollte man nun Gleichung und Ungleichung kombinieren, so wird die Regularitätsbedingung wahrscheinlich wesentlich schärfer und würde lineare Unabhängigkeit der Gradienten aller wesentlich aktiven Funktionen fordern. Problematisch ist auch, dass bei zwei PC^1 -Funktionen sehr schwer klar wird, welche Indizes man wirklich als insgesamt aktiv zählt, damit der Gradient auf jeden Fall enthalten ist.

Für zwei oder mehr Ungleichungen lässt sich diese Frage noch einigermaßen zufrieden stellend klären. Seien $F_1, \dots, F_k : \mathbb{R}^n \rightarrow \mathbb{R}$ PC^1 -Funktionen, dann gilt

$$M = \{x | F_1(x) \leq 0, \dots, F_k(x) \leq 0\} = \left\{ x \mid \max_{i=1, \dots, k} F_i(x) \leq 0 \right\},$$

wodurch wir das Problem auf eine sinnvolle neue PC^1 -Funktion zurückgeführt haben.

Für zwei Gleichungen mit F, G würde man durch eine separate Betrachtung nur schwer Zugang zu den neuen wesentlichen aktiven Indizes erhalten. Übergang zum Quadrat oder Betrag für eine neue gemeinsame PC^1 -Funktion würde die Regularität verletzen. Die wesentlich aktiven Indizes von $F - G$ haben die Gestalt „ ij “ zu $f_i - g_j$. Mit diesen neuen Indizes betrachte man nun

$$M = \{x | F(x) = 0, G(x) = 0\} = \{x | F(x) - G(x) = 0\} \cap \{x | G(x) = 0\}.$$

Dann sieht man leicht, dass lineare Unabhängigkeit der verschiedenen $Df_i(x), Dg_j(x)$ hinreichend für die Regularitätsbedingung für Normalenkegel bei Schnitt von Mengen ist. Allerdings für den Erhalt einer Gleichung ist noch eine wesentlich aufwendigere Untersuchung der Mengen und der wesentlich aktiven Indizes nötig.

Für eine Mischung von Gleichung und Ungleichung kann eventuell eine ähnliche Untersuchung helfen.

Folgerung 10.38 (Anwendung auf den Graphen) Es gilt $\text{gph } F = \{(y, t) | F(y) - t = 0\}$. Mit $M_2 := \text{gph } F$ folgt nun ohne zusätzliche Bedingungen:

$$\mathbb{R}(\{Df_i(x) | i \in I_F^w(x)\} \times \{-1\}) \cup N_{\text{gph } F}^F(x) \subset N_{\text{gph } F}^{\text{lim}, F}(x) \subset N_{\text{gph } F}^{\text{Cl}}(x) \text{ und}$$

$$N_{\text{gph } F}^{\text{Cl}}(x) = \text{conv}(\mathbb{R} \{(Df_i(x), -1) | i \in I_F^w(x)\}).$$

Beweisanmerkung: Die Ableitungen erhalten (wegen der zusätzlichen Ableitung nach t in der Gleichung) eine „-1“, in der letzten Komponente. Dadurch ist die Regularitätsbedingung immer erfüllt, da bei Konvexkombinationen diese Komponente gleich bleibt und nie 0 wird. \square

Diese Formel zeigt nun, warum die Berechnung des „Clarke-Co-Differentials“ im eindimensionalen Fall sinnlos ist, denn hier wird bei PC^1 -Funktionen mit nicht verschwindender Ableitung immer \mathbb{R} raus kommen und sonst hat man hier auch kaum brauchbare Aussagen.

Folgerung 10.39 (einfache Optimalitätsbedingung) Seien $G, F : \mathbb{R}^n \rightarrow \mathbb{R}$ PC^1 -Funktionen mit zugehörigen Funktionen g_j, f_i um x .

Wie zuvor definieren wir $M_1 = \{y | F(y) \leq 0\}$ bzw. $M_2 = \{y | F(y) = 0\}$ und fordern wieder $0 \notin \text{conv} \{Df_i(x) | i \in I_F^w(x)\}$ und $0 \notin \text{conv} \{Dg_j(x) | j \in I_G^w(x)\}$.

Sei x ein lokal minimaler Punkt von G auf M_1 bzw. M_2 , dann gilt

$$\exists y \in \mathbb{R}_+^l, z \in \mathbb{R}_+^m \text{ bzw. } \mathbb{R}^m, \sum_{j \in I_G^w(x)} y_j = 1 : 0 = \sum_{j \in I_G^w(x)} y_j Dg_j(x) + \sum_{i \in I_F^w(x)} z_i Df_i(x).$$

Statt der wesentlich aktiven Indizes kann man auch die einfach zu ermittelten aktiven Indizes verwenden, welche ja gleich sind bei guter Wahl der g_j, f_i .

Beweis: Die Bedingung folgt aus den berechneten Formeln, wenn man sie in die Optimalitätsbedingung bzgl. des Clarke-Konzepts einsetzt. \square

In der letzten Bemerkung haben wir uns größere Gedanken über die wesentlich aktiven Indizes gemacht, wenn wir mehrere Restriktionen verknüpfen. Nimmt man darauf keine Rücksicht und sucht nur nach Optimalitätsbedingungen, so kommt man mit Hilfe von [Cla83] schnell zu nichtglatten Karush-Kuhn-Tucker Optimalitätsbedingungen unter einer nichtglatten Version der MFCQ. Hierbei ist gleich anzumerken, dass man natürlich auch calmness hätte verwenden können, die in [Cla83] bevorzugt wird, allerdings lässt sich MFCQ leichter prüfen und passt deshalb besser zu unseren konkreten Berechnungen.

Satz 10.40 (allgemeine Optimalitätsbedingung durch „MFCQ“)

Seien $F : \mathbb{R}^n \rightarrow \mathbb{R}$, $G : \mathbb{R}^n \rightarrow \mathbb{R}^m$ und $H : \mathbb{R}^n \rightarrow \mathbb{R}^l$ PC^1 -Funktionen um x mit zugehörigen stetig differenzierbaren Funktionen $f_{i_F}, g_{i_G}^j$ zu G_j und $h_{i_H}^j$ zu H_j , mit $i_F = 1, \dots, p; i_G = 1, \dots, q; i_H = 1, \dots, r$. Es sei

$$M = \{x | G_i(x) \leq 0 \forall i \in \{1, \dots, m\}, H_j(x) = 0 \forall j \in \{1, \dots, l\}\}$$

und x ein lokaler Minimalpunkt von F auf M . Dann existieren $\alpha \geq 0, y \in \mathbb{R}_+^m, z \in \mathbb{R}^l$ mit $(\alpha, y, z) \neq 0$, sodass gilt

$$0 \in \alpha \partial^{Cl} F(x) + \sum_{j=1}^m y_j \partial^{Cl} G_j(x) + \sum_{j=1}^l v_j \partial^{Cl} H_j(x) \text{ und } \langle y, G(x) \rangle = 0. \quad (10.8)$$

Folglich existieren $a \in \mathbb{R}^p, B \in \mathbb{R}_+^{q \times m}, C \in \mathbb{R}^{r \times l}$ mit $(a, B, C) \neq 0, a_i = 0 \forall i \notin I_F^w, b_{ij} = 0 \forall i \notin I_{G_j}^w, b_{ij} = 0 \forall i \notin I_{G_j}^w$ und

$$0 = \sum_{i=1}^p a_i Df_i + \sum_{(i,j)=1}^{(q,m)} b_{ij} Dg_i^j(x) + \sum_{(i,j)=1}^{(r,l)} c_{ij} Dh_i^j(x).$$

Die Komplementaritätsbedingung lautet dann $b_{ij} = 0 \forall j : G_j(x) < 0$.

Existiert nun außerdem eine „MFCQ-Richtung“ $u_0 \in \mathbb{R}^n$ mit

$$Dg_i^j(x)u_0 < 0 \forall i \in I_{G_j}^w(x) \forall j : G_j(x) = 0 \text{ und } Dh_i^j(x)u_0 = 0 \forall i \in I_{H_j}^w(x), j = 1, \dots, l$$

und sind die $Dh_i^j(x)$ mit $i \in I_{H_j}^w(x)$ und $j = 1, \dots, l$ linear unabhängig, so gilt $\alpha \neq 0$ und wir können oBdA $\alpha = 1$ bzw. $\|a_i\| = 1$ wählen.

Beweis: Der erste Teil ist eine Anwendung von Theorem 6.1.1 in [Cla83] und unserer Rechenformel für das Subdifferential. Der Beweis dieses Theorems ist recht aufwendig und braucht sehr viele Hilfssätze, da man wegen der mehrelementigen Subdifferenziale nicht mehr die lineare Dualität sofort anwenden kann.

Da a, B, C nur für das Produkt der Faktoren der Konvexkombination in den Subdifferenzialen der PC^1 -Funktionen mit α, y bzw. z folgt ihre Existenz sofort und es gilt $a = 0, B = 0$ bzw. $C = 0$ genau dann, wenn $\alpha = 0, y = 0$ bzw. $z = 0$.

Angenommen also $\alpha = 0$ bzw. $a = 0$. Nun bilden wir das Skalarprodukt der Gleichung (10.8) mit u_0 und erhalten

$$0 = 0 + \sum_{(j:G_j(x)=0)} \sum_{(i \in I_{G_j}^w(x))} b_{ij} Dg_i^j(x)u_0 + 0 < 0 \text{ nach „MFCQ“, falls } B \neq 0.$$

Folglich gilt $B = 0$ und wir erhalten die Gleichung

$$0 = \sum_{j=1}^l \sum_{(j \in I_{H_j}^w(x))} c_{ij} Dh_i^j(x).$$

Aufgrund der linearen Unabhängigkeit der wesentlich aktiven $Dh_i^j(x)$ folgt sofort $C = 0$, was im Widerspruch zu $(a, B, C) \neq 0$ steht. Folglich gilt $\alpha > 0$ bzw. $a \in \mathbb{R}_+^p - \{0\}$. \square

Eine andere Bedingung erhalten wir mithilfe von [KS94]. Allerdings ist dazu eine genaue Kenntnis der Richtungsableitung von PC^1 -Funktionen nötig. Diese ist stückweise linear bzgl. der Ableitungsrichtung, da sie lipschitzstetig ist und eine Auswahl aus den Ableitungen, der zugrundeliegenden C^1 -Funktionen ist, die natürlich linear sind bzgl. der Ableitungsrichtung. Somit können wir wieder Richtungsableitungen und Clarke-Subdifferential bestimmen. Die Bedingungen für die (in diesem Fall mehreren) MFCQ-Richtungen in [KS94] beziehen sich dann auf diese Ableitungen. Ich nenne hier eine abgeschwächte aber dafür einfachere Version, damit die Bedingung verständlicher und vor allem besser nachprüfbar ist.

Definition 10.41 (PC^1 -MFCQ) Wir betrachten eine Menge $M \subset \mathbb{R}^n$ mit $x \in M$, welche beschrieben ist durch PC^1 -Funktionen G, H wie oben. Sei G eine stetige Auswahl von C^1 -Funktionen g_1, \dots, g_m . H sei eine Auswahl von C^1 -Funktionen h_1, \dots, h_l . Diese können in manchen Komponenten lokal um x identisch sein. Die unterschiedlichen Komponenten bezeichnen wir mit q_1, \dots, q_p . OBdA gelte $G(x) = 0$.

Zum einen fordern wir nun, dass zu jeder Auswahl von höchstens n Funktionen der q_j die zugehörigen Gradienten linear unabhängig sind. Zum anderen fordern wir, dass zu jedem $v \in \mathbb{R}^n$ mit $H'(x; v) = 0$ eine Richtung $u(v) \in \mathbb{R}^n$ existiert mit $Dg_j(x)u(v) < 0 \forall j \in I_G^w(x)$ und $Dh_i(x)u(v) = 0 \forall i \in I_H^w(x) : Dh_i(x)v = 0$.

Bemerkung (Kommentare zur Vereinfachung der ursprünglichen Bedingung) Die zusätzliche Bedingung $Dh_i(x)v = 0$ für die $i \in I_H^w(x)$ kommt daher, dass ich nur die linearen Teile der Richtungsableitung h betrachten will, die in Richtung v aktiv sind. Hier werden eventuell mehr Indizes betrachtet, als die wesentlich aktiven, aber diese stärkere Bedingung wird wesentlich einfacher zu untersuchen.

Fordert man zusätzlich $u(v) = v$, so sollte die Bedingung im Allgemeinen leicht nachzuprüfen sein, allerdings ist dieser Zwischenschritt über die w recht elegant, da in Richtung v eines der q_i in mehreren Komponenten aktiv sein kann und man mit dem w im weitesten Sinne eines auswählt und den Rest ignoriert.

Statt der $j \in I_G^w(x)$ wurden ursprünglich nur die Indizes betrachtet, die bei der Richtungsableitung nach v noch aktiv sind.

Ist diese Definition sehr kompliziert und schwer nachzuprüfen, so ist der zugehörige Satz zur Optimalität sehr einfach. Er kommt wie die vorige Definition aus [KS94]. Der Beweis ist sehr aufwendig, verteilt über den ganzen Artikel und benutzt im wesentlichen einen richtungsdifferenzierbaren Homeomorphismus in einer Umgebung von x zur Umwandlung des Problems.

Satz 10.42 (Optimalität mit PC^1 -MFCQ) Erfülle obige Restriktionsmenge M die PC^1 -MFCQ in einem lokal optimalen x zur Ziel- PC^1 -Funktion F . Dann gilt

$$F'(x; v) \geq 0 \forall v : H'(x; v) = 0 \wedge G'_i(x; v) \leq 0 \forall i \in I_G^w(x).$$

Bemerkung (Berechnungsprobleme) Die hier erwähnten Formeln vermitteln zwar das Gefühl, dass man mit PC^1 -Funktionen noch relativ gut umgehen kann, aber sie verschweigen auch einiges.

Erstens kann allein das Ermitteln der wesentlich aktiven Indexfunktion aufwendig sein.

Zweitens haben wir zwar eine Formel für die Richtungsableitung, aber in ihr steckt ein Grenzwert, der auszurechnen ist. Ein „Ausweg“ ist hier, dass man durch die wesentlich aktiven Indizes eine Einschränkung der möglichen Ableitungen hat. Dadurch wäre beim praktischen Umgang zwar ein Grenzwertverfahren mit einem gewissen Fehler nötig, aber letztendlich würde man sich für einen Index entscheiden, den man für die Richtungsableitung wählt und hätte eine meist exakte Näherung.

Der dritte Punkt ist der gravierendste. Wenn man nur weiß, dass die Funktion eine PC^1 -Funktion ist, aber weder Indexfunktion noch die einzelnen Funktionen kennt, aus denen

die Funktion zusammengesetzt ist, so kann man zwar das Wissen über die entsprechenden Formeln nutzen, aber das Prüfen von Regularität ist nicht mehr möglich und auch konkrete Berechnungen von Optima werden schwieriger bzw. sind höchstens näherungsweise möglich.

10.5 ...zur Normalenkegelabbildung zum positiven Orthanten

In diesem Abschnitt berechnen wir den Limiting-Fréchet-Normalenkegel an den Graphen der Normalenkegelabbildung zum positiven Orthanten.

Wir wissen bereits

$$N_{\mathbb{R}_+^n}^\alpha(x) \stackrel{(6.1)}{=} \prod_{j=1}^n \begin{cases} \{0\}, & \text{falls } x_j < 0 \\ \mathbb{R}, & \text{falls } x_j = 0 \end{cases}, \forall \alpha.$$

Für $G(x) = N_{\mathbb{R}_+^n}^{lim,F}(x)$ kann man nun $N_{\text{Gph } G}^{lim,F}(x, y)$ untersuchen mit $y \in N_{\mathbb{R}_+^n}^{lim,F}(x)$. Dies wird in [Out06] getan in Lemma 3.4 und liefert die Formel

$$N_{\text{Gph } G}^{lim,F}(x, x^*) = \prod_{j=1}^n M_j(x_j, y_j) \prod_{j=1}^n N_j(x_j, y_j) \text{ mit}$$

$$M_j(x_j, y_j) \times N_j(x_j, y_j) = \begin{cases} 0 \times \mathbb{R}, & \text{falls } x_j > 0 \\ \mathbb{R} \times 0, & \text{falls } x_j = 0, y_j < 0 \\ (0 \times \mathbb{R}) \cup (\mathbb{R} \times 0) \cup (\mathbb{R}_- \times \mathbb{R}_+), & \text{falls } x_j = y_j = 0 \end{cases}$$

11 Mehrstufenoptimierung

Beim Lösen von Optimierungsaufgaben kann es passieren, dass die Restriktionmenge „zusätzlich“ eingeschränkt wird durch die Forderung, dass bestimmte Variablen aus anderen Optimierungsaufgaben kommen. Dies kann dann wieder auch für die andere Optimierungsaufgabe gelten, wodurch mehrere Stufen entstehen. Das Hauptproblem ist dabei, dass die notwendige Optimalitätsbedingung für das Problem der unteren Stufe im Allgemeinen nicht differenzierbar ist. Somit muss diese extra behandelt werden. Zwei Ansätze dazu werden wir im folgenden betrachten und ein wenig vergleichen.

Hierbei wird gleichzeitig eine andere Möglichkeit zusätzlicher Restriktionen modelliert. Diese Art wird Gleichgewichtsbedingung genannt. Diese kann oft über Optimierungsaufgaben modelliert werden aber auch durch normale Restriktionen und eine Komplementaritätsbedingung.

In den zugehörigen Abkürzungen steht „CC“ für „Komplementaritätsbedingung“, „EC“ für „Gleichgewichtsbedingung“, „G“ für „verallgemeinert“ und „MP“ für „mehrstufige Optimierungsaufgabe“.

11.1 GMPEC

Zunächst betrachten wir den Modellierungsansatz von J. Outrata [Out06].

Definition 11.1 (GMPEC) Sei $F : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ stetig differenzierbar und $Q : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ abgeschlossen. Die Zielfunktion $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ sei zumindest lokal lipschitzstetig.

Wir betrachten die Aufgabe

$$\begin{aligned} & \text{minimiere } f(x, y) \\ & \text{mit} \\ & \quad 0 \in F(x, y) + G(y) \\ & \quad (x, y) \in N. \end{aligned}$$

$0 \in F(x, y) + G(y)$ ist hier die Gleichgewichtsbedingung. Die zulässige Menge übersetzen wir in die Form von Satz 8.1 mittels

$$g(x, y) = (y, -F(x, y)) \text{ und } Q = \text{Gph}(G).$$

Somit lautet die Regularitätsbedingung (7.3):

Definition 11.2 (OMFCQ)

$$\left. \begin{aligned} & \begin{pmatrix} 0 & E \\ -D_x F(x, y) & -D_y F(x, y) \end{pmatrix}^T \begin{pmatrix} u \\ v \end{pmatrix} \in -N_N^{\text{lim}, F}(x, y) \\ & (u, v) \in N_{\text{Gph}(G)}^{\text{lim}, F}(y, -F(x, y)) \end{aligned} \right\} \Rightarrow (u, v) = 0.$$

Folgerung 11.3 (Normalenkegel) Nun folgt für $M = \{(x, y) \in N \mid 0 \in F(x, y) + G(y)\}$ mit dem Satz 8.1: $N_M^{\text{lim}, F} \subset \dots$

$$\dots \subset \left\{ \begin{pmatrix} D_x F(x, y)v + a \\ D^* G(y, -F(x, y))(v) + D_y F(x, y)v + b \end{pmatrix} \middle| v \in \mathbb{R}^m, (a, b) \in N_N^{\text{lim}, F}(x, y) \right\},$$

wobei hier das obige (u, v) durch das Co-Differential umschrieben wurde, denn $(u, -v) \in N_{\text{Gph}(G)}^{\text{lim}, F}(y, -F(x, y))$ ist äquivalent zu $u \in D^* G(y, -F(x, y))(v)$.

Folgerung 11.4 Sei nun (x', y') ein lokal optimaler Punkt und f um diesen Punkt lipschitzstetig.

Wenn dieser Punkt die Regularitätsbedingung erfüllt, dann folgt mit der Optimalitätsbedingung aus Satz 9.4: $\exists (a', b') \in N_N^{\text{lim}, F}(x', y'), (c', d') \in \partial^{\text{lim}, F} f(x', y'), v' \in \mathbb{R}^m$

$$\begin{aligned} 0 &= c' + D_x F(x', y')v' + a' \\ 0 &\in d' + D_y F(x', y')v' + b' + D^* G(y', -F(x', y'))(v'). \end{aligned}$$

11.2 GMPCC

Ein etwas anderes konkreteres Modell ist jenes in [SS00] zu MPCC's. Hierbei wird mit verallgemeinerten Komplementaritätsbedingungen (Complementarity Constraints) modelliert.

Definition 11.5 (GMPCC) Seien die matrixwertige Komplementaritätsfunktion $F : \mathbb{R}^n \rightarrow \mathbb{R}^{m \times l}$, die Restriktionsfunktionen $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ und $h : \mathbb{R}^n \rightarrow \mathbb{R}^q$ und die Zielfunktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ stetig differenzierbar.

Wir betrachten die Aufgabe

$$\begin{aligned} & \text{minimiere } f(x) \\ & \text{mit} \\ & \quad \min \{F_{k1}(x), \dots, F_{kl}(x)\} = 0, \quad k = 1, \dots, m \\ & \quad g(x) \leq 0, \\ & \quad h(x) = 0. \end{aligned}$$

Natürlich gilt $m \geq 1$ und $l \geq 2$. Hierbei verallgemeinert man die Standardkomplementaritätsbedingung der nichtlinearen parametrische Optimierung

$$F_1(x, y) \geq 0, \quad F_2(x, y) = x \geq 0, \quad \langle F_1(x, y), x \rangle = 0.$$

Für diese Aufgabe gibt es mehrere Regularitätsbedingungen, welche in [SS00] nachzulesen sind. Ich beschränke mich hier auf die MFCQ-Bedingung, da diese für folgende Betrachtungen ausreicht.

Definition 11.6 (CCMFCQ) Eine Aufgabe erfüllt diese Bedingung in x , wenn die Gradienten

$$DF_{ki}(x) \text{ mit } F_{ki}(x) = 0 \text{ und } Dh_s(x) \text{ mit } s = 1, \dots, q$$

linear unabhängig sind und es eine Richtung v gibt mit

$$Dg_r(x)v < 0 \quad \forall r : g_r(x) = 0 \text{ und } DF_{ki}(x)v = 0 \quad \forall (k, i) : F_{ki}(x) = 0, Dh_s(x)v = 0.$$

Satz 11.7 (Optimalitätsbedingung) Sei x ein lokales Minimum des GMPCC's und gelte die CCMFCQ in x . Dann existieren $\Gamma \in \mathbb{R}^{m \times l}$, $y \in \mathbb{R}_+^p$ und $z \in \mathbb{R}^q$ mit

$$\begin{aligned} 0 &= Df(x)^T - DF(x)^T \Gamma + Dg(x)^T y + Dh^T(x)z, \\ 0 &= F_{ki}(x) \Gamma_{ki} \quad \forall (k, i), \\ 0 &\leq \Gamma_{ki} \Gamma_{kj} \quad \forall (i, j, k) : F_{ki}(x) = F_{kj}(x) = 0, \\ 0 &= \langle g(x), y \rangle. \end{aligned}$$

Anmerkung: Der Beweis in [SS00] ist ähnlich zum klassischen. Allerdings benutzt er eine verallgemeinerte Lagrange-Bedingung, welche in [Cla83] bewiesen wird, und das Clarke-Subdifferential der Funktion $\min \{F_{k1}(x), \dots, F_{kl}(x)\}$, dessen Formel im Abschnitt zu PC^1 -Funktionen zu finden ist. \square

11.3 Vergleich von GMPEC und GMPCC

Damit wir mit der Notation nicht durcheinander kommen, werden die Größen des GMPEC mit einem zusätzlichen \prime bezeichnen. Zunächst „verallgemeinern“ wir das GMPEC

und lassen auch $G'(y')$ von x' abhängig sein. Sei also $x = (x', y')$. Somit vereinfacht sich die Regularitätsbedingung zu

$$\left. \begin{array}{l} -u + DF'(x)^T v \in N_{N'}^{lim,F}(x) \\ (u, v) \in N_{\text{Gph}(G')}^{lim,F}(x, -F'(x)) \end{array} \right\} \Rightarrow (u, v) = 0.$$

Bemerkung (Grundannahmen für das GMPEC) Für die Beziehung $0 \in F'(x) + G'(x)$ nehmen wir an, dass sie unter Zuhilfenahme zusätzlicher Variablen sich in Gleichungen, Ungleichungen und Komplementaritätsbedingungen schreiben lässt.

Das heißt, wir haben $F' = 0$, $G'_k(x) = \min \{F_{k1}(x), \dots, F_{kl}(x)\} = 0$, $k = 1, \dots, m$ und $\bar{N} = \{x \mid \bar{g}(x) \leq 0, \bar{h}(x) = 0\}$. Somit ist $G' =: F$ nicht mehr mehrwertig, aber im Allgemeinen nicht differenzierbar.

Außerdem fordern wir, dass sich N' in Form von Gleichungen und Ungleichungen schreiben lässt und die Bedingungen aus \bar{N} auch enthält: $N' = \{x \mid g(x) \leq 0, h(x) = 0\}$. Letztendlich fordern wir noch, dass N' regulär ist, d.h. MFCQ bzgl. h und g ist erfüllt.

Somit gilt

$$N_{N'}^\alpha(x) = \left\{ x^* \mid x^* = \sum_{i=1}^p Dg_i(x)y_i + \sum_{j=1}^q Dh_j(x)z_j, y \in \mathbb{R}^p, y \geq 0, z \in \mathbb{R}^q \right\}.$$

$$\text{Es gilt } (u, v) \in N_{\text{Gph}(G')}^{lim,F}(x, 0) \Leftrightarrow u \in D^{lim,F}G'(x, 0)(-v).$$

Wir haben bereits gesehen, dass man für $D^{lim,F}F(x, 0)(-v)$ im Allgemeinen keine konkrete Formel erwarten kann. Deshalb verschärfen wir $(u, v) \in N_{\text{Gph}(G')}^{lim,F}(x, 0)$ in der Regularitätsbedingung zu

$$u \in \text{conv}(D^{lim,F}G'(x, 0)(-v)).$$

Gehen wir wieder davon aus, dass alle aktiven Indizes wesentlich aktiv sind, so erhalten wir

$$u \in \left\{ -\sum_{k=1}^m v_k \sum_{i=1}^l \lambda_{ki} DF_{ki}(x) \mid \min \{F_{ki}(x), \lambda_{ki}\} = 0, \sum_{i=1}^l \lambda_{ki} = 1 \ \forall k \right\}$$

mithilfe der Formel zum Co-Differential von PC^1 -Funktionen, wobei hier die Ableitungen Gradienten sind. Da $F' \equiv 0$ und $v \in \mathbb{R}^m$ beliebig ist, ersetzen wir v und λ_{ki} durch eine neue Variable α_{ki} und erhalten

$$-u = \sum_{k=1}^m \sum_{i=1}^l \alpha_{ki} DF_{ki}(x).$$

Wir fordern nur noch $\alpha_{ki} F_{ki}(x) = 0$ und haben den Zusammenhang $v = 0 \Leftrightarrow \alpha = 0$. Dieser gilt allerdings nur, wenn die Gradienten der aktiven $F_{ki}(x)$ nicht verschwinden. Eine leichte Überlegung zeigt, dass diese Bedingung in der OMFCQ enthalten ist ($y = 0$, $z = 0$, entsprechendes $\lambda_{ki} = 1$, $v = e_k$). Also fordert die Regularitätsbedingung nun:

Für alle $y \in \mathbb{R}_+^p$ mit $y_i g_i(x) = 0 \ \forall i$, $\alpha \in \mathbb{R}^{m \cdot l}$ mit $\alpha_{ki} F_{ki}(x) = 0 \ \forall k, i$ und $z \in \mathbb{R}^q$ mit

$$0 = \sum_{k=1}^m \sum_{i=1}^l \alpha_{ki} DF_{ki}(x) + \sum_{i=1}^p Dg_i(x)y_i + \sum_{j=1}^q Dh_j(x)z_j \text{ folgt } (\alpha, y, z) = 0,$$

In Satz 8.8 haben wir aber bereits gezeigt, dass diese Bedingung äquivalent umgeformt werden kann in eine MFCQ-Bedingung. Diese ist in diesem Fall die CCMFCQ.

Geht man also von „regulären“ Mengen und Funktionen aus, so sind die beiden Bedingungen OMFCQ und CCMFCQ äquivalent.

Durch die Regularität fallen die verallgemeinerten Normalenkegel zusammen. Gehen wir nun von einer stetig differenzierbaren Zielfunktion aus, so fallen auch die verallgemeinerten Subdifferenziale zusammen. Dadurch vereinfacht sich die Optimalitätsbedingung zum GMPEC und man kann zeigen, dass sie der Optimalitätsbedingung für das GMPCC entspricht.

12 Formelsammlung

In diesem Abschnitt sollen alle gefundenen Gleichungen wiederholt werden zum einfachen Nachschlagen. Ungleichungen und Teilmengenbeziehungen werden nicht wiederholt. Damit alles noch übersichtlich bleibt und auch den Charakter einer Formelsammlung hat, verzichte ich auf Text und Kommentare und treffe folgende Vereinbarungen:

Für bestimmte Details und Beweise kann man in der Diplomarbeit nachschlagen. Zusätzliche Voraussetzungen sind erwähnt und Zwischendefinitionen, die für bestimmte Formeln nötig sind, mit „:=“ gekennzeichnet.

Generell sind alle nicht mehrwertigen Abbildungen lokal um x lipschitzstetig. Abbildungen bilden immer zwischen den passenden Räumen ab. Dabei werden Kleinbuchstaben für Funktionale verwendet. Mengen sind prinzipiell abgeschlossen und von der Dimension her passend zur Formel.

Satz 12.1 (Normal- und Tangentialkegel)

$$\begin{aligned} N_M^{Cl}(x) &= \text{conv}(N_M^{lim,F}(x)) \\ N_M^F(x) &= N_M^B(x) \\ N_M^F(x) &= N_M^B(x) = N_M^{lim,F}(x) = N_M^{Cl}(x), \text{ falls } M \text{ konvex} \\ T_M^B(x) &= T_M^{Cl}(x), \text{ falls } M \text{ konvex} \\ M &:= N \times Q \\ N_M^\alpha(x, y) &= N_N^\alpha(x) \times N_Q^\alpha(y) \text{ mit } \alpha \in \{F; lim, F; Cl\} \end{aligned}$$

$$\begin{aligned} M &:= \{x | h(x) = 0, g(x) \leq 0\} \text{ erfülle MFCQ} \\ N_M^\alpha(x) &:= N_M^F(x) = N_M^B(x) = N_M^{lim,F}(x) = N_M^{Cl}(x) \\ T_M^B(x) &= T_M^{Cl}(x) = \{u | Dh(x)u = 0, Dg_i(x)u \leq 0 \forall i : g_i(x) = 0\} \\ N_M^\alpha(x) &= \left\{ \sum_{i=1}^p Dh_i(x)u_i + \sum_{j=1}^q Dg_j(x)v_j \mid u \in \mathbb{R}^q, v \in \mathbb{R}_+^p, v_j = 0 \forall j : g_j(x) < 0 \right\} \end{aligned}$$

$$\begin{aligned}
N_M^{\alpha, \epsilon}(x) &:= \{v \mid \langle v, y - x \rangle \leq \epsilon \|y - x\| \forall y \in M \text{ nahe } x\} \\
N_M^{\text{lim}, F}(x) &= \limsup_{x_k \xrightarrow{M} x, \epsilon_k \searrow 0} N_M^{\alpha, \epsilon_k}(x_k) = \limsup_{x_k \xrightarrow{M} x} N_M^F(x_k) \\
M &:= N \cap G^{-1}(Q) \text{ mit differenzierbarem } G \\
N_M^\alpha(x) &= DG(x)^T N_Q^\alpha(G(x)) + N_N^\alpha(x) \quad \forall \alpha \neq ko, \text{ falls} \\
&\quad N \text{ und } Q \text{ regulär und GMFCQ erfüllt, siehe Satz 8.1} \\
M &:= N \cap Q \\
N_M^\alpha(x) &= N_Q^\alpha(x) + N_N^\alpha(x) \quad \forall \alpha \neq ko, \text{ falls} \\
&\quad N \text{ und } Q \text{ regulär und } N_Q^\alpha(x) \cap -N_N^\alpha(x) = \{0\}, \text{ siehe Satz 8.4}
\end{aligned}$$

Satz 12.2 (Subdifferential vs. Normalenkegel)

$$\begin{aligned}
\partial^\alpha f(x) &= \left\{ x^* \mid (x^*, -1) \in N_{\text{epi } f}^\alpha(x, f(x)) \right\} \\
\partial^\alpha i_M(x) &= N_M^\alpha(x) \\
M &:= \{x \mid f(x) \leq 0\} \\
N_M^\alpha(x) &= \mathbb{R}_+ \partial^\alpha f(x), \quad \forall \alpha \neq ko, \text{ falls } 0 \notin \partial^{Cl} f(x), f(x) = 0, \partial^{Cl} f(x) = \partial^F f(x)
\end{aligned}$$

Satz 12.3 (Subdifferential) alternative Definitionen:

$$\begin{aligned}
\partial^F f(x) &= \{v \in X \mid \langle v, y - x \rangle \leq f(y) - f(x) + o(y - x) \quad \forall y \in X\} \\
\partial^{\text{lim}, F} f(x) &= \left\{ v \in X \mid \exists x_k \rightarrow x, v_k \rightarrow v : v_k \in \partial^F f(x_k) \right\} \\
\partial^{Cl}(f)(x) &= \text{conv} \{v = \lim Df(x_k) \mid x_k \rightarrow x, Df(x_k) \text{ existiert}\}
\end{aligned}$$

Gleichungen:

$$\begin{aligned}
\partial^F f(x) &= \partial^{\text{lim}, F} f(x) = \partial^{Cl}(f)(x) = \{Df(x)\}, \text{ falls } f \text{ um } x \text{ stetig differenzierbar} \\
\partial^F f(x) &= \partial^{\text{lim}, F} f(x) = \partial^{Cl}(f)(x) = \partial^{ko} f(x), \text{ falls } f \text{ konvex} \\
f^{lD}(x; u) &:= \inf \{c \mid c \in Cf(x; u)\} \text{ Definition der „lower Dini derivative“} \\
\partial^F f(x) &= \partial^{ko} f^{lD}(x; \cdot)(0) \\
\partial^{Cl}(f)(x) &= \partial^{ko}(f^{Cl}(x; \cdot))(0) \\
\partial^{Cl}(f)(x) &= \text{conv}(\partial^{\text{lim}, F} f(x))
\end{aligned}$$

Satz 12.4 (andere verallgemeinerte Ableitungen)

$$\begin{aligned}
CF(x; u) &= \{F'(x; u)\}, \text{ falls } F'(x; u) \text{ existiert} \\
CF(x; u) &= TF(x; u) = DF(x)u, \text{ falls } F \text{ um } x \text{ stetig differenzierbar} \\
f^{Cl}(x; u) &= \sup \{v | v \in Tf(x; u)\} \\
\text{conv}(TF(x; u)) &= \partial^{Cl}F(x)(u), \text{ verallg. Clarke-Ableitung hier benutzt} \\
Tf(x; u) &= \partial^{Cl}f(x)(u) \\
D^\alpha F(x, F(x))(v) &= \partial^\alpha (\langle v, F(\cdot) \rangle)(x), \alpha \in \{F; \text{lim}, F\} \\
D^{\text{lim}, F}(x, f(x))(1) &= \partial^{\text{lim}, F}f(x) \\
D^F F(x, F(x))y &= \{DF(x)^T y\}, \text{ falls } F \text{ in } x \text{ differenzierbar} \\
D^{\text{lim}, F} F(x, F(x))y &= \{DF(x)^T y\}, \text{ falls } F \text{ um } x \text{ stetig differenzierbar}
\end{aligned}$$

$\text{conv}(D^{\text{lim}, F} F(x, F(x))(v)) = \{A^T v | A \in \partial^{Cl}(F)(x)\}$, mit verallg. Clarke-Ableitung siehe S.405 [RW98]

Satz 12.5 (Umkehrabbildung)

$$\begin{aligned}
C(F^{-1}) &= (CF)^{-1} \\
T(F^{-1}) &= (TF)^{-1} \\
D^\alpha(F^{-1}) &= (D^\alpha F)^{-1} \forall \alpha
\end{aligned}$$

Satz 12.6 (Summenformeln)

$$\begin{aligned}
T(F + G)(x; u) &= TF(x; u) + TG(x; u) \\
C(F + G)(x; u) &= CF(x; u) + CG(x; u) \\
T(F)(x; u + v) &= TF(x; u) + TF(x; v) \\
\partial^F(f + g)(x) &= \partial^F(f)(x) + \partial^F(g)(x) \\
\partial^{\text{lim}, F}(f + g)(x) &= \partial^{\text{lim}, F}(f)(x) + \partial^{\text{lim}, F}(g)(x) \\
\partial^{Cl}(f + g)(x) &= \partial^{Cl}(f)(x) + \partial^{Cl}(g)(x)
\end{aligned}$$

Dies gilt, wenn f bzw. F um x differenzierbar ist. Bei den Subdifferentialen haben wir außerdem Gleichheit, falls f und g konvex sind, vom Maximumtyp sind zumindest in einer Umgebung von x oder wenn aus einem anderen Grund beide Funktionen regulär sind.

$$\partial^\alpha \left(\sum_{i=1}^m f_i(x_i) \right) (x_1, \dots, x_m) = \partial^\alpha f_1(x_1) \times \dots \times \partial^\alpha f_m(x_m) \forall \alpha$$

Sei G hier mehrwertig und F in x differenzierbar.

$$\begin{aligned}
D^F(F + G)(x, y)(y^*) &= DF(x)^T y^* + D^F G(x, y - F(x))(y^*) \\
D^{\text{lim}, F}(F + G)(x, y)(y^*) &= DF(x)^T y^* + D^{\text{lim}, F} G(x, y - F(x))(y^*), \\
&\text{falls } F \text{ zusätzlich um } x \text{ stetig differenzierbar}
\end{aligned}$$

Satz 12.7 (positive Skalierbarkeit) $\lambda \in \mathbb{R}, \lambda > 0$

$$\begin{aligned} CF(x, y; \lambda u) &= \lambda CF(x, y; u) \\ TF(x, y; \lambda u) &= \lambda TF(x, y; u) \\ C(\lambda F)(x, y; u) &= \lambda CF(x, y; u) \\ T(\lambda F)(x, y; u) &= \lambda TF(x, y; u) \\ \partial^\alpha(\lambda f)(x) &= \lambda \partial^\alpha(f)(x) \quad \forall \alpha \end{aligned}$$

Satz 12.8 (Kettenregeln)

$$\begin{aligned} F : \mathbb{R}^n \times \mathbb{R}^m &\rightarrow \mathbb{R}^q, F(x, y) = h(x, g(y)) \\ h : \mathbb{R}^n \times \mathbb{R}^p &\rightarrow \mathbb{R}^q \text{ in } p\text{-Komponente lipschitzstetig differenzierbar} \\ g : \mathbb{R}^m &\rightarrow \mathbb{R}^p \text{ einfach bzw. richtungsdifferenzierbar} \end{aligned}$$

Dann gilt:

$$\begin{aligned} TF((x, y); (u, v)) &= T_x h((x, g(y)); u) + D_g h(x, g(y)) Tg(x; v), \\ CF((x, y); (u, v)) &= C_x h((x, g(y)); u) + D_g h(x, g(y)) Cg(x; v). \end{aligned}$$

$$\begin{aligned} G : \mathbb{R}^n &\rightarrow \mathbb{R}^m \\ H : \mathbb{R}^m &\rightarrow \mathbb{R}^p \\ f : \mathbb{R}^n &\rightarrow \mathbb{R}^p, f := H(G(\cdot)) \end{aligned}$$

Dann gilt:

$$\begin{aligned} Cf(x; u) &= CH(G(x); CG(x; u)), \text{ falls } H \text{ oder } G \text{ richtungsdifferenzierbar} \\ Tf(x; u) &= DH(G(x))(TG(x; u)), \text{ falls } H \text{ um } x \text{ stetig differenzierbar} \end{aligned}$$

Folgerung 12.9 (Produktregel) Seien $M : \mathbb{R}^n \rightarrow \mathbb{R}^{a \times b}$ und $N : \mathbb{R}^q \rightarrow \mathbb{R}^{b \times c}$ matrixwertige Funktionen und eine von beiden *einfach* bzw. richtungsdifferenzierbar. Für $F : \mathbb{R}^{n+q} \rightarrow \mathbb{R}^{a \times c}$, $F(x, y) := M(x)N(y)$ erhalten wir dann

$$\begin{aligned} TF((x, y); (u, v)) &= TM(x; u)N(y) + M(x)TN(y; v), \\ CF((x, y); (u, v)) &= CM(x; u)N(y) + M(x)CN(y; v) \end{aligned}$$

Satz 12.10 (Ableitungen von PC^1 -Funktionen) Seien f und F PC^1 -Funktionen mit Komponenten f_i .

$$\begin{aligned} \partial^\alpha f(x) &= \text{conv} \{ Df_i(x) | F(x) = f_i(x) \} \alpha \neq ko \text{ f\u00fcr Maximumtypfunktionen} \\ \partial^{Cl} F(x) &= \text{conv} \{ Df_i(x) | i \in I_F^w(x) \} \\ F'(x; u) &= Df_i(x)u \in \{ Df_i(x)u | i \in I_F^w(x), \exists t_k \searrow 0, i \in I_F^w(x + tu) \} \\ CF(x; u) &= \{ F'(x, u) \} \\ \partial^F f(\cdot)(x) &= \partial^{ko} f'(x; \cdot)(0) \\ \text{conv}(TF(x; u)) &= \text{conv}(\{ Df_i(x)u | i \in I_F^w(x) \}) \\ Tf(x; u) &= \text{conv}(\{ Df_i(x)u | i \in I_F^w(x) \}) \\ \text{conv}(D^{lim, F} F(x, F(x))(v)) &= \text{conv} \{ Df_i(x)^T v | i \in I_F^w(x) \} \end{aligned}$$

Bemerkung Hat man Gleichungen für Normalenkegel bzw. Tangentialkegel und verallgemeinerte Ableitungen, so resultiert dies in konkreteren Optimalitätsbedingungen, wie zum Beispiel die KKT-Bedingungen.

Satz 12.11 (Kojima)

$$\Phi(x, y, z) := \begin{pmatrix} \Phi^1(x, y, z) \\ \Phi^2(x, y, z) \\ \Phi^3(x, y, z) \end{pmatrix} : \mathbb{R}^{n+p+q} \rightarrow \mathbb{R}^{n+p+q} \text{ mit}$$

$$\Phi^1(x, y, z) := (Df(x)^T + Dg(x)^T y^+ + Dh(x)^T z)^T$$

$$\Phi^2(x, y, z) := g(x) - y^-$$

$$\Phi^3(x, y, z) := h(x)$$

$$x \text{ ist optimalverd\"achting} \Leftrightarrow (x, y^+, z) \text{ ist KKT-Punkt} \Leftrightarrow \Phi(x, y^+ + g(x)) = 0$$

Es wurden in Satz 10.16 auch die Kontingent-Ableitung und die Thibault-Ableitung der Kojima-Funktion berechnet.

Satz 12.12 (stationäre Punkte) In Satz 10.20 gezeigt:

$$u \in CX((p, x); v) \Leftrightarrow$$

$$v \in DM(x)uN(y', z') + M(x)CN((y', z'); \mathbb{R}^{p+q}) \text{ f\"ur ein } (y', z') : \Phi(x, y', z') = p$$

Weitere Gleichungen liefern Optimalitätsbedingungen für PC^1 -Funktionen oder MPCC's und der Limiting-Fréchet-Normalenkegel an die Normalenkegelabbildung an den positiven Orthanten.

13 Thesen

- In diesem Abschnitt möchte ich meine persönliche Sicht auf die gewonnenen Erkenntnisse in dieser Arbeit zusammenfassen.
- Zunächst kann man festhalten, dass es zahlreiche Zusammenhänge zwischen den verallgemeinerten Ableitungen gibt. Die Kontingent-Ableitung entspricht in den schönen Fällen der Richtungsableitung. Das Fréchet-Subdifferential entsteht dann, indem man das Subdifferential der Richtungsableitung bildet. Über Bildung von Grenzwerten bzw. konvexer Hülle erhält man jeweils die anderen betrachteten Subdifferentialie. Für lipschitzstetige Funktionen erhält man dadurch dann wieder Aussagen für die Co-Differentialie. Darüber und über Grenzwertbildung bei der Kontingent-Ableitung kommt man letztendlich zu einer Teilmenge der Thibault-Ableitung und kann Aussagen über deren konvexe Hülle machen. Für Abbildungen in den \mathbb{R}^m und mehrwertige Abbildungen gibt es ähnliche Zusammenhänge, allerdings keine Subdifferentialie.
- Durch die Bildung der Grenzwerte, Übergang zur konvexen Hülle oder Wechsel von festen Ableitungsrichtungen zu Grenzwerten von Richtungen können die jeweiligen Ableitungsbegriffe sehr stark auseinander fallen. Dies gilt selbst bei relativ einfachen Funktionen, wie im Beispielabschnitt zu sehen war.
- Rechenregeln liefern meist nur bei Differenzierbarkeit oder komplizierteren Forderungen Gleichungen und ansonsten nur Teilmengenbeziehungen. Will man Gleichheit haben, so fordert man verschiedenste Formen der „Regularität“, welche dann zur Folge haben, dass die Ableitungen oder zugehörige Normalenkegel gleich sind bzw. die Bildung der Grenzwerte nichts ändert. Damit kann man die verschiedenen Rechenregeln kombinieren und erhält somit Gleichheit, womit man das Problem der Berechnung der verallgemeinerten Ableitungen auf einfachere Probleme reduziert ist, deren Lösung man kennt.
- Stellt man sich nun die große Frage: „Was ist denn nun eigentlich berechenbar?“, so muss man zuerst schauen, wie können Mengen und Funktionen überhaupt gegeben sein.
- In dem Fall, dass man differenzierbare Funktionen verwendet, die durch Polynome, Sinus, Brüche usw. aufgebaut sind, sind die verallgemeinerten Ableitungen geradezu uninteressant, da durch die bekannten Ketten- und Ableitungsregeln die klassischen Ableitungen berechenbar sind und man daraus die verallgemeinerten schlussfolgern kann. Zwar gibt es in diesem Bereich auch Herausforderungen, aber diese rechtfertigen nicht die verallgemeinerten Ableitungen.
- Hat man Mengen beschrieben durch eine Gleichung, so ist schon eine Regularitätsbedingung nötig um ad hoc eine Formel angeben zu können. Ohne diese Bedingung ist in vielen Fällen eine Berechnung der Normalenkegel noch möglich, indem man Fallunterscheidungen in den Punkten, welche die Bedingung verletzen,

durchführt. Das berühmteste Beispiel ist hier wohl die Gleichung $xy = 0$ und die zugehörigen Normalenkegel im Nullpunkt. Betrachtet man zu dieser Menge Optimalitätsbedingungen, so stellt man fest, dass selbst bei differenzierbarer Zielfunktion Grenzwertbildung oder anschließende konvexe Hülle wesentlich schwächere Bedingungen bzw. gar keine Aussage liefern.

- Sind die betrachteten Funktionen nun nicht mehr von solch einfacher Gestalt, so kann die Ableitung verschwinden oder sehr komplex werden. In vielen Fällen, kann man jedoch von PC^1 -Funktionen ausgehen und diese sind noch richtungsdifferenzierbar. Hierbei ist zu beachten, dass diese Eigenschaft zwar günstig sein kann für Algorithmen, aber leider im Allgemeinen keine konkrete Formel liefert sondern nur eine Gesetzmäßigkeit. Aus meiner Sicht zeigt dieser Funktionstyp, warum verallgemeinerte Ableitungen nützlich sein können. Man kann zur Funktion eine konkrete Formel für das Clarke-Subdifferential angeben und sofort Optimalitätsbedingungen angeben. Natürlich sind diese nun im Allgemeinen wesentlich schwächer, da die entsprechende Ableitung sehr groß ist. Interessanterweise ergaben sich durch die Berechnung des Clarke-Subdifferentials noch andere konkrete Formeln. Hier habe ich schnell erkannt, dass dieses Thema sehr komplex wird, wenn man die Restriktionsmenge mit PC^1 -Funktionen beschreibt. Für spezielle Funktionen, die zur Komplementarität führen, gibt es zwar schon Untersuchungen, aber ansonsten scheint mir das ein Forschungsgebiet zu sein, wo es noch einiges zu entdecken gibt.
- Funktionen zu betrachten, die komplexer sind als PC^1 -Funktionen, scheint für mich aus der Sicht der Berechenbarkeit nicht sinnvoll zu sein. Die Blitzfunktion, die Sägezahntypfunktionen und die Umkehrfunktion zu x^3 zeigen, dass diese Funktionen zum Teil schwer beschreibbar sind und eine Berechnung der Ableitungen nur über Umwege möglich ist.
- Dahingegen sind mehrstufige Optimierungsaufgaben schon etwas besser untersucht und einfacher zu betrachten. Die beiden von mir betrachteten Konzepte sind beides Verallgemeinerungen des klassischen Komplementaritätsproblems, dass durch die KKT-Bedingung entsteht. Für zumindest eine konkret betrachtete Aufgabe liefern sie identische Bedingungen. Das Modell, welches S. Scholtes verwendet [SS00], liefert sofort einfach überprüfbare Regularitäts- und Optimalitätsbedingungen. Ein anderes, welches J. Outrata [Out06] verwendet, benutzt eine Regularitätsbedingung, welche wesentlich aufwendiger zu prüfen ist und meist weitere Regularitätsvoraussetzungen benötigt. Ein großer Vorteil ist, dass sie auch auf allgemeinere Aufgaben anwendbar ist als nur die klassische Restriktion mit Komplementaritätsbedingung. Mir ist allerdings noch nicht klar, wozu so etwas praktisch nötig ist und wann die komplexeren Nebenbedingungen nicht aus anderen Optimierungsaufgaben oder Komplementaritätsbedingungen kommen.

14 Symbolverzeichnis

\mathbb{R}^q	der q -dimensionale reelle Vektorraum
\mathbb{R}_+^q	der positive Orthant des Vektorraums
\mathbb{R}_-^q	der negative Orthant des Vektorraums
$\mathbb{R}^{p \times q}$	Menge der $p \times q$ -Matrizen
$T_M^B(x)$	Kontingenz-Kegel an M in x
$T_M^{Cl}(x)$	Clarke-Tangentialkegel an M in x
$N_M^{ko}(x)$	Normalenkegel der konv. Opt. an M in x
$N_M^F(x)$	Fréchet-Normalenkegel an M in x
$N_M^B(x)$	Bouligand-Normalenkegel an M in x
$N_M^{lim,F}(x)$	Limiting-Fréchet-Normalenkegel an M in x
$N_M^{Cl}(x)$	Clarke-Normalenkegel an M in x
$\partial^\alpha(x)$	α -Subdifferential von f in x
$\text{int}(M)$	Menge der inneren Punkte von M
\overline{M}	Abschluss von M
$\text{Rand}(M)$	Menge der Randpunkte von M
$\text{conv } M$	konvexe Hülle von M
$CF(z; u)$	Kontingenz-Ableitung von F im Punkt z in Richtung u
$TF(z; u)$	Thibault-Ableitung von F im Punkt z in Richtung u
$f'(x; u)$	Ableitung von f im Punkt x in (positive) Richtung u
$D^\alpha(z)v$	α -Co-Differential im Punkt z in Richtung v
$DF(x)$	klassische Ableitung
$\text{sign}(x)$	Vorzeichenfunktion: $\text{sign}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases}$
$i_M(x)$	Indikatorfunktion: $i_M(x) = \begin{cases} 0 & , x \in M \\ \infty & , x \notin M \end{cases}$
PC^1	stückweise differenzierbar
$I_F(x)$	aktive Indizes einer PC^1 -Funktion F in x
$I_F^w(x)$	wesentlich aktive Indizes einer PC^1 -Funktion F in x

15 Literatur

- [Cla83] F.H. Clarke, Optimization and Nonsmooth Analysis, J. Wiley& Sons, 1983
- [KK02] D. Klatte und B. Kummer, Nonsmooth equations in Optimization, Kluwer Academic Publishers, 2002
- [Kum88] B. Kummer, Newton's method for non-differentiable functions, Advances in Mathematical Optimization, Akademie Verlag, 1988
- [KS94] L. Kuntz und S. Scholtes, A nonsmooth variant of the Mangasarian-Fromovitz constraint qualification, Journal of Optimization Theory and Applications 82, 1994
- [Kum08] B. Kummer, Optimization and Variational Inequalities, elektronische Zusammenfassung, 2008
- [KK05] D. Klatte and B. Kummer, Strong Lipschitz Stability of Stationary Solutions for Nonlinear Programs and Variational Inequalities, SIAM J. Optimization, 2005, 16: 96–119
- [Mor06] B.S. Mordukhovich, Variational Analysis And Generalized Differentiation I, Springer-Verlag, 2006
- [Out06] J.V. Outrata, Mathematical programs with equilibrium constraints: Theory and numerical methods, Nonsmooth Mechanics of solids, 2006
- [Roc03] R.T. Rockafellar, A property of piecewise smooth functions, Comput. Optim. Appl. 25, 2003, 247-250
- [RW98] R.T. Rockafellar und R. Wets, Variational Analysis, Springer-Verlag, 1998
- [SS00] H. Scheel und S. Scholtes, Mathematical program with complementarity constraints: Stationarity, optimality and sensitivity, Mathematics of Operations Research 25, 2000
- [Sch94] S. Scholtes, Introduction to piecewise differentiable equations, Institut für Statistik und Mathematische Wirtschaftstheorie Universität Karlsruhe, 1994

16 Erklärungen

Selbstständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Berlin, den

Einverständniserklärung

Hiermit erkläre ich mich damit einverstanden, dass ein Exemplar meiner Diplomarbeit in die Bibliothek des Instituts für Mathematik verbleibt.

Berlin, den