

Bachelorarbeit

Proximal-Punkt Algorithmen auf
Riemannschen Mannigfaltigkeiten

Alexandra Pruhs
Matrikelnummer 405503

Münster, 29.09.2022

Inhaltsverzeichnis

1. Einleitung	1
2. Darstellung des Optimierungsproblems	3
3. Grundlagen der Riemannschen Geometrie, der konvexen Analysis und Regularisierung	4
3.1 Einführung in relevante Themen der Riemannschen Geometrie	4
3.2.1 Einführung des Konvexitäts-Begriffs von Mengen	11
3.2.2 Einführung in das Gebiet der konvexen Funktionen	16
3.3 Regularisierung	19
3.3.1 Die Moreau-Yosida Regularisierung	20
3.3.2 Wohldefiniertheit der Moreau-Yosida Regularisierung	21
4. Der Proximal-Punkt Algorithmus	24
4.1 Einführung des Proximal-Punkt Algorithmus	24
4.2 Wohldefiniertheit des Proximal-Punkt Algorithmus	25
4.3. Konvergenz	26
5. Zusammenfassung	32
Bild- und LiteraturvZusammenfassungverzeichnis	

Kapitel 1

Einleitung

Aufbauend auf dem Bachelorseminar "Optimierung und Variationsmethoden in der Bildverarbeitung und den Naturwissenschaften", welches am Institut für Numerische und Angewandte Mathematik der Westfälischen Wilhelms- Universität von Prof. Dr. Benedikt Wirth gehalten wurde, behandelt diese Bachelorarbeit die Thematik der Proximal-Punkt-Algorithmen auf Riemannschen Mannigfaltigkeiten und beinhaltet einen Einblick in die zentralen Konzepte des Themas.

Das Forschungsgebiet der Optimierungsmethoden beschäftigt sich mit der optimalen Steuerung von biologischen, chemischen, physikalischen und ökonomischen Prozessen. Beispielhaft für diese Themengebiete zu nennen sind zum einen inverse Probleme in der Medizin und Biologie, zum anderen aber auch das Design von optimalen Bauteilen für Ingenieurwissenschaften. Ferner werden Optimierungsmethoden in der Bildverarbeitung genutzt, um unter anderem Streuungsprobleme und Bildsegmentierungen zu behandeln. Hier ist ein spezielleres Beispiel die Bildwiederherstellung, wobei gestörte Bilder durch Approximation dem Originalbild wieder angeglichen werden können. Die Optimierung von Bildern findet unter anderem Anwendung in der Medizin bei den bildgebenden Verfahren wie zum Beispiel der Computertomographie.

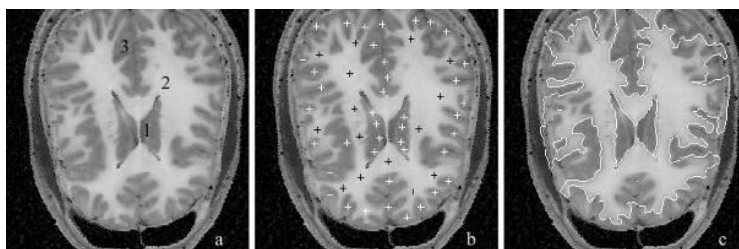


Abb: Segmentierung in der medizinischen Bildverarbeitung (Quelle 12)

Diesem Bachelorarbeit eröffnet den Einblick in das Konzept der Proximal Punkt-Algorithmen auf Riemannschen Mannigfaltigkeiten, mit dessen Hilfe in der Optimierung ein Minimierungsproblem gelöst wird. Hierbei handelt es sich um eine Einschränkung des

euklidischen Raums auf das Gebiet einer Riemannschen Mannigfaltigkeit, welches betrachtet werden soll. Um die Grundlagen für diese Betrachtung zu schaffen, wird zunächst auf die Eigenschaften von Mannigfaltigkeiten Bezug genommen. Die Gebiete der Riemannschen Geometrie sowie der konvexen Analysis bilden die Grundlage für dieses Verfahren.

Nach den einleitenden Worten zum Thema der Optimierung, bietet das folgende Kapitel zunächst einen Einblick in die einführenden Themen. Hier wird vorab das übergeordnete Optimierungsproblem dargestellt. Darauf folgend wird mit der Einführung von grundlegenden mathematischen Aspekten der riemannschen Geometrie begonnen. Dabei wird zunächst ein Überblick über die für das Thema relevanten Definitionen und Eigenschaften von Mannigfaltigkeiten geschaffen. Im folgenden Abschnitt wird näher auf die Konvexität von Mengen und Funktionen eingegangen.

Nachdem alle grundlegenden Themen eingeführt wurden, gehen wir dann zum Bereich der Regularisierung über, um anschließend den Proximal-Punkt-Algorithmus zu betrachten. In Bezug auf diesen wollen wir uns dann mit der Wohldefiniertheit und der Konvergenz dessen beschäftigen.

Kapitel 2

Darstellung des Optimierungsproblem

Der Inhalt dieses Unterkapitels orientiert sich an [1, S.1 f.]

Im Folgenden wird zunächst das Modellproblem für das Proximal-Punkt-Verfahren eingeführt. Bei dem Modellproblem handelt es sich im Speziellen um ein Minimierungsproblem.

Dazu sei M eine vollständige riemannsche Mannigfaltigkeit und $f: M \rightarrow \mathfrak{R}$ eine konvexe Funktion. Darauf bezogen betrachten wir nun folgendes Problem:

$$\min_{x \in M} f(x)$$

Ferner sei $p_0 \in M$ ein Startpunkt und $\{\lambda_k\}$ eine Folge positiver Zahlen. Dann erzeugt das Proximal-Punkt Verfahren auf der riemannschen Mannigfaltigkeit genau diese Folge $\{x_k\} \subset M$ durch folgende Iteration:

$$x_{k+1} = \operatorname{argmin}_{y \in M} \{f(y) + \lambda_k p_{x(k)}(y)\}$$

Dabei ist $p_x: M \rightarrow \mathfrak{R}$ gegeben durch $p_x(y) = \frac{1}{2} d^2(x, y)$, wobei $d(x, y)$ die riemannsche Distanz darstellt.

Ziel dieses Optimierungsproblems ist es zu zeigen, dass die Erweiterung der bereits auf dem euklidischen Raum bekannten Konzepte ebenfalls auf dem eingeschränkten Gebiet der Riemannschen Mannigfaltigkeit mit nicht positiver Schnittkrümmung natürlich ist. Hierfür wird im Verlauf dieser Ausfertigung die Wohldefiniertheit der durch

$$x_{k+1} = \operatorname{argmin}_{y \in M} \{f(y) + \lambda_k p_{x(k)}(y)\}$$

generierten Folge gezeigt, sowie die Konvergenz der Folge gegen ein Minimum für den Fall, dass M eine Riemannsche Mannigfaltigkeit mit nicht positiven Schnittkrümmungen ist.

Kapitel 3

Grundlagen der Riemannschen Geometrie, der konvexen Analysis und Regularisierung

Der Inhalt dieses Kapitels orientiert sich an [1, S.1 f.] , [2, S.1 f.], [3, S.1 f.] , [4, S.1 f.], [5, S.22f.],[6, S.1 f.], [7 S.1 f.] , [8, S.1 f.] und [9, S.95f.].

Nachdem wir das Modellproblem im vorherigen Kapitel dargestellt haben, folgt nun in diesem Kapitel die Darstellung der grundlegenden Konzepte der Riemannschen Geometrie und der konvexen Analysis.

Hierbei werden wir einen genaueren Blick auf den Raum der Riemannschen Mannigfaltigkeiten werfen, sowie deren Eigenschaften. Ebenso betrachte wir konvexe Mengen und konvexe Funktionen und stellen hier einen Zusammenhang her.

Abschließend werden wir in diesem Kapitel auf Basis der zuvor definierten Aussagen auf die Regularisierung, welche uns dabei hilft, das Verfahren zu stabilisieren.

3.1 Einführung in relevante Themen der Riemannschen Geometrie

Die Riemannsche Geometrie ist das Gebiet der Differentialgeometrie, welches sich zur Aufgabe gemacht hat, Riemannsche Mannigfaltigkeiten zu untersuchen. Sie wurde erstmals im 19. Jahrhundert vom Mathematiker Bernhard Riemann aufgestellt. Dabei hat er sich mit dem Bereich der Geometrie beschäftigt, in welchem sich metrische Eigenschaften von Punkt zu Punkt verändern können. Daraus haben sich insbesondere lokale Betrachtungen ergeben, welche unter anderem Kurven und Winkel mit einbeziehen. Um uns einen Überblick über die Aufgaben der Riemannschen Geometrie zu verschaffen, wollen wir in diesem Unterkapitel die grundlegenden Themen der Riemannschen Geometrie kurz erläutern. Dabei werden wir alle Aspekte, welche von Relevanz sind, für den weiteren Verlauf dieser Ausarbeitung besprechen. Hierfür starten wir nun mit der Definition der Riemannschen Mannigfaltigkeit.

Definition 3.1

Sei M eine differenzierbare n -dimensionale Mannigfaltigkeit und g ein Tensorfeld, welches jedem Punkt $p \in M$ eine euklidische Metrik des Tangentialraums $T_p M$ zuordnet. Dann heißt (M, g) Riemannsche Mannigfaltigkeit und g Riemannsche Metrik.

Die Riemannsche Metrik g auf der Mannigfaltigkeit M sollte die Form einer positiv definiten, symmetrischen Bilinearform haben. Diese ist differenzierbar von p abhängig und lässt sich wie folgt darstellen: $g_p : T_p M \times T_p M \rightarrow \mathfrak{R}$. Für den Fall, dass g_p nicht positiv definit ist, sprechen wir für (M, g) von einer Pseudo-Riemannschen Mannigfaltigkeit.

Ein einfaches Beispiel für eine Riemannsche Mannigfaltigkeit ist der euklidische Raum \mathfrak{R}^n , welcher auf natürliche Weise mit der Riemannschen Metrik versehen ist und wie folgt dargestellt wird: $\langle \cdot, \cdot \rangle$. Mit Hilfe der Riemannschen Metrik lassen sich die wesentlichen geometrischen Eigenschaften von Mannigfaltigkeiten beschreiben. Sie beschreibt nicht eine Metrik im eigentlichen Sinne, sondern ist vielmehr ein Skalarprodukt.

Die Riemannsche Metrik bietet die Möglichkeit, Längen und Winkel von Tangentialvektoren zu messen, wie zum Beispiel den Winkel zwischen zwei sich schneidenden Kurven. Ebenso lässt sich die Länge der Kurven durch Integration bestimmen und das Längenfunktional wird wie folgt dargestellt:

$$L(c) := \int_a^b \sqrt{g_{c(t)}(c'(t), c'(t))} dt$$

mit $c: [a, b] \rightarrow M$ und (M, g) ist Riemannsche Mannigfaltigkeit.

Minimiert man dieses Längenfunktional auf dem Gebiet aller Kurven, so erhält man die Riemannsche Distanz $d(x, x')$.

Ein weiterer Punkt, mit dem sich die Differentialgeometrie im Zusammenhang mit Mannigfaltigkeiten beschäftigt, ist die Krümmung von Kurven. Hier spielen vor allem die Geodäten, welche die kürzeste Verbindung zwischen zwei Punkten darstellen, eine wichtige Rolle. Wenn man die Krümmung einer Fläche im euklidischen Raum betrachtet, ist bekannt, dass diese relativ einfach durch die zweite Ableitung zu bestimmen ist. Dagegen gestaltet sich dies bei der Betrachtung von Riemannschen Mannigfaltigkeiten etwas schwieriger. Die Tatsache, dass bei differenzierbaren Mannigfaltigkeiten nicht einfach der Differenzenquotient für die zweite Ableitung einer Kurve gebildet werden kann, besteht aus dem Grund, dass sich die zwei Vektoren in unterschiedlichen Vektorräumen befinden. Um dieses Problem zu lösen, wird eine Abbildung definiert, welche einen Zusammenhang zwischen diesen Vektorräumen herstellt. Einer dieser Zusammenhänge ist der Levi-Civita-Zusammenhang, mit dessen Hilfe die Richtungsänderung eines Vektorfeldes in die Richtung eines weiteren Vektorfeldes bestimmt werden kann.

Dies gleicht der Parallelverschiebung eines Tangentialvektors entlang glatter Kurven in einer Mannigfaltigkeit und keiner globalen Parallelverschiebung.

Wir beschreiben den Levi-Civita-Zusammenhang nun wie folgt:

Satz 3.2

Seien X, Y zwei Vektorfelder auf dem euklidischen Raum \mathbb{R}^n mit $Y = \sum_{i=1}^n f_i \frac{\partial}{\partial x_i}$. Dann beschreibt der Levi-Civita-Zusammenhang die Richtungsableitung von Y nach X , wobei diese durch $\nabla_X Y = Y = \sum_{i=1}^n \nabla_X f_i \frac{\partial}{\partial x_i}$ beschrieben wird.

Sei dann (M, g) eine Riemannsche Mannigfaltigkeit, dann wissen wir, dass genau ein Zusammenhang ∇ dieser Art existiert, welcher metrisch und torsionsfrei ist. Dieser wird als Levi-Civita-Zusammenhang bezeichnet und ist durch die folgende Gleichung gegeben:

$$2g(\nabla_X Y, Z) = Xg(Y, Z) + Yg(Z, X) - Zg(X, Y) + g([X, Y], Z) - g([Y, Z], X) + g([Z, X], Y)$$

Im Folgenden nennen wir einige weitere Fakten:

Bemerkungen 3.3

1. Ein Vektorfeld X entlang einer Kurve c ist genau dann parallel, wenn $\nabla_c X = 0$
2. Wenn c' selbst parallel ist, dann heißt c geodätisch.
3. Die geodätische Gleichung $\nabla_{\gamma'} \gamma' = 0$ ist eine nichtlineare gewöhnliche Differentialgleichung zweiter Ordnung mit $\gamma = \gamma_v(\cdot, x)$, wobei x die deren Position bestimmt und v die Geschwindigkeit in x darstellt.
4. Wenn $\|\gamma'\| = 1$ dann heißt γ normalisiert.

Nachdem wir im vorherigen Abschnitt bereit den Begriff der Geodäte kurz angesprochen haben, wollen wir nun noch eine etwas genauere Definition hierfür geben:

Definition 3.4

Ein geodätisches Segment ist die Einschränkung einer Geodäte auf ein abgeschlossenes und beschränktes Intervall.

Ein solches geodätisches Segment, welches zwei Punkte x und x' auf der Mannigfaltigkeit M verbindet heißt minimal, wenn die Länge der riemannschen Distanz $d(x', x)$ entspricht. Dann werden die Geodäte auch minimierende Geodäte genannt.

In den vorangegangenen Abschnitten konnten wir nun die Themen einführen, welche wir nun für die Darstellung der Vollständigkeit von Riemannschen Mannigfaltigkeiten benötigen.

Satz 3.5

Eine zusammenhängende riemannsche Mannigfaltigkeit (M, g) heißt geodätisch vollständig, wenn alle Geodäten auf ganz \mathfrak{X} definiert sind.

Geodätisch vollständig bedeutet also nichts anderes, als das eine Geodäte γ für alle Punkte $p \in M$ und für alle Tangentialvektoren $v \in T_p M$ auf \mathfrak{X} definiert ist.

Eine relevante Rolle im Zusammenhang mit der Vollständigkeit von riemannschen Mannigfaltigkeiten spielt das Theorem von Hopf-Rinow, wessen Hauptaussage ist, dass bei Riemannschen Mannigfaltigkeiten der Begriff der geodätischen Vollständigkeit gleichzusetzen ist mit der Vollständigkeit im Sinne metrischer Räume.

Theorem 3.6: Hopf-Rinow

Sei (M, g) eine endl. dimensionale zusammenhängende Riemannsche Mannigfaltigkeit. Dann sind folgende Aussagen äquivalent und es lassen sich zwei Punkte durch eine Geodäte verbinden:

1. (M, g) ist geodätisch vollständig.
2. (M, g) ist vollständig als metrischer Raum
3. Für den metrischen Raum (M, g) gilt die Heine-Borel-Eigenschaft d.h. das jede abgeschlossene und beschränkte Teilmenge kompakt ist
4. "Die Exponentialabbildung hat maximalen Definitionsbereich, d.h. eine und damit jede der folgenden Aussagen gelten:
 - a) exp ist auf ganz $T_p M$ definiert.
 - b) Für jedes $p \in M$ ist exp_p auf ganz $T_p M$ definiert.
 - c) Es gibt einen Punkt $p \in M$, sodass exp_p auf ganz $T_p M$ definiert ist."

(5, S.23)

Da der Beweis dieser Aussage recht umfangreich ist, verzichten wir hier auf die Beweisführung.

Für alle weiteren Zusammenhänge werden wir im Folgenden von der Vollständigkeit der verwendeten Mannigfaltigkeiten ausgehen.

Ein weiterer Begriff, welcher relevant sein wird für die folgenden Kapitel ist der Begriff der exponentiellen Karte. Diese stellt die natürlichste Form einer Retraktionsabbildung auf Riemannschen Mannigfaltigkeiten dar. Die exponentielle Karte wird in der Riemannschen Geometrie beschrieben als eine Karte einer Teilmenge des Exponentialraums $T_x M$ einer Riemannschen Mannigfaltigkeit M zur Mannigfaltigkeit selbst. Dabei wird die exponentielle Karte in Form der folgenden Abbildung $exp_x: T_x M \rightarrow M$ dargestellt und wird folgendermaßen definiert: $exp_x v = \gamma_v(1, x)$.

Ein weiteres Thema, auf welches wir eingehen wollen, sind die Begriffe des Krümmungstensors und der Schnittkrümmung. Diese benötigen wir in erster Linie für die Definition der sogenannten Hadamard-Mannigfaltigkeit. Diese wiederum spielt im weiteren Verlauf eine tragende Rolle. Daher wollen wir auch hier kurz auf die Begrifflichkeiten eingehen und die wichtigsten Punkte kurz definieren:

Definition 3.7

Sei M eine riemannsche Mannigfaltigkeit und seien X, Y, Z Vektorfelder auf M . Dann definieren wir den Krümmungstensor R wie folgt:

$$R(X, Y) = \nabla_x \nabla_y Z - \nabla_y \nabla_x Z - \nabla_{[Y, X]} Z$$

mit $[X, Y] = YX - XY$.

Die aus dem Krümmungstensor resultierende Schnittkrümmung ist dann wie folgt definiert:

$$K(X, Y) = \langle R(X, Y)Y, X \rangle / (\|X\|^2 \|Y\|^2 - \langle X, Y \rangle^2)$$

wobei $\|X\|^2 = \langle X, X \rangle$ ist.

Falls $K(X, Y) \leq 0$ für alle X und Y , dann wird M eine riemannsche Mannigfaltigkeit von nicht-positiven Schnittkrümmungen genannt.

Die Schnittkrümmung kann als Verallgemeinerung der Gaußschen Krümmung gesehen werden und ordnet jeder Ebene im Tangentialraum einer riemannschen Mannigfaltigkeit eine bestimmte Krümmung zu. Diese Krümmung wird durch den Krümmungstensor beschrieben. Dieser Riemannsche Krümmungstensor, welcher nach dem Mathematiker Bernhard Riemann benannt wurde, wird benötigt um darzustellen, wie ähnlich eine riemannsche Mannigfaltigkeit lokal betrachtet dem \mathbb{R}^n ist.

Fasst man die letzten Erkenntnisse zusammen, so kann man die sogenannte Hadamard-Mannigfaltigkeit bereits kurz und knapp definieren:

Definition 3.8

Eine vollständige, einfach zusammenhängende Mannigfaltigkeit mit nicht positiver Schnittkrümmung nennt man dann Hadamard-Mannigfaltigkeit.

Die soeben definierte Hadamard-Mannigfaltigkeit kann im Vergleich zu anderen Formen der Mannigfaltigkeit eine Besonderheit aufweisen, die im nächsten Theorem beschrieben wird.

Theorem 3.9

Sei M eine Hadamard-Mannigfaltigkeit. Dann ist M diffeomorph zum euklidischen Raum \mathbb{R}^n .
Genauer heißt dies, dass die exponentielle Karte $\exp_x: T_x M \rightarrow M$ an jedem Punkt $x \in M$ ein Diffeomorphismus ist.

Da Diffeomorphismen bijektive, stetig differenzierbare Abbildungen sind, deren Umkehrabbildungen ebenfalls stetig differenzierbar sind ist die logische Konsequenz dieses Theorems, dass die Differential Struktur und die topologischen Eigenschaften der Hadamard-Mannigfaltigkeit dieselben sein müssen wie im euklidischen Raum.

Eine weitere Definition, welche wir im Verlauf dieser Ausarbeitung benötigen, ist die des geodätischen Dreiecks, welches als Dreieck auf einer gekrümmten Fläche, wie zum Beispiel auf der Erdoberfläche beschrieben werden kann. Die folgende Definition liefert uns die Bezeichnungen und Zusammenhänge dieses besonderen Dreiecks.

Definition 3.10

Sei M eine riemannsche Mannigfaltigkeit. Dann bezeichnet $\Delta(x_1, x_2, x_3)$ ein geodätisches Dreieck, welches aus drei verschiedenen Punkten (Ecken) und drei minimierenden geodätischen Segmenten γ_{i+1} (Seiten) besteht. Dabei verbinden die Seiten γ_{i+1} die Punkte x_{i+1} und x_{i+2} .

Die Besonderheit von Dreiecken auf Mannigfaltigkeiten besteht darin, dass deren Winkelsumme nicht genau 180° betragen muss, sondern dass diese durchaus auch größer sein kann als 180° .

Daraus folgt folgendes Theorem, welches seine Anwendung in den folgenden Kapiteln findet, in denen es um konvexe Mengen geht. Auch hier wird das zuvor definierte geodätische Dreieck verwendet und ein Zusammenhang zwischen den Seiten und Winkeln hergestellt. Es findet seine Anwendung in der Beweisführung einiger relevanter Theoreme.

Theorem 3.11

Sei M eine Hadamard Mannigfaltigkeit. Sei das geodätische Dreieck $\Delta(x_1, x_2, x_3)$ gegeben, wie soeben definiert. Bezeichne $\gamma_{i+1}: [0, l_{i+1}] \rightarrow M$ ein geodätisches Segment, welches die Punkte x_{i+1} und x_{i+2} miteinander verbindet und definiere die Seiten mit $l_{i+1} := L(\gamma_{i+1})$ und die Winkel durch $\theta_{i+1} := \angle(\gamma'_{i+1}(0) - \gamma'_i(l_i))$, mit $i = 1, 2, 3 \pmod{3}$.

Dann gelten folgende Aussagen: 1) $\theta_1 + \theta_2 + \theta_3 \leq \pi$

$$2) l_{i+1}^2 + l_{i+2}^2 - 2l_{i+1}l_{i+2}\cos\theta_{i+2} \leq l_i^2$$

$$3) l_{i+1}\cos\theta_{i+2} + l_i\cos\theta_i \geq l_{i+2}$$

Den Beweis zu diesem Theorem ist in [10] zu finden.

3.2 Einführung in die Konvexe Analysis

Im folgenden Unterkapitel werden fundamentale Eigenschaften der konvexen Analysis auf Mannigfaltigkeiten eingeführt, welche im späteren Verlauf dieser Ausarbeitung benötigt werden.

Der Themenbereich der Konvexität spielt in den verschiedensten Bereichen der Mathematik eine Rolle. Besonders in den verschiedensten Themengebieten der Funktionalanalysis aber auch in der Optimierungstheorie, sowie in weiteren anwendungsbezogenen Teilgebieten, findet sie ihre Anwendung. In der Differentialgeometrie werden insbesondere Punkte, Kurven und Flächen sowie Untermannigfaltigkeiten betrachtet, welche durch differenzierbare Funktionen dargestellt werden können.

In diesem Kapitel wird es im Folgenden um die Geometrie konvexer Mengen und Funktionen auf dem Gebiet der Riemannschen Mannigfaltigkeiten gehen, im speziellen auf Hadamard-Mannigfaltigkeiten.

Wir werden zunächst auf den Begriff der Konvexen Menge und deren Eigenschaften im Allgemeinen eingehen und im Anschluss auf die Konvexität auf Hadamard-Mannigfaltigkeiten.

3.2.1 Einführung des Konvexitäts-Begriffs von Mengen

In diesem Unterkapitel werden wir uns zunächst mit den allgemeinen Definitionen der Konvexität auf Mengen beschäftigen. Daraufhin folgen spezifische Definitionen der Konvexität auf Mannigfaltigkeiten, sowie die Darstellung einiger Eigenschaften dieser Mengen. Die zwei wichtigsten Themen in diesem Unterkapitel sind jedoch die Existenz einer eindeutigen Projektion sowie die Existenz eines Supports.

Definition 3.12

Sei die Menge $K \subset \mathfrak{R}^n$ und seien $x, y \in K$. Dann heißt die Menge K konvex genau dann, wenn mit $x, y \in K$ auch deren Verbindungsstrecke in K enthalten ist. Diese Verbindungsstrecke ist wie folgt definiert: $[x, y] := \{\lambda x + (1 - \lambda)y : 0 \leq \lambda \leq 1\} \in K$.

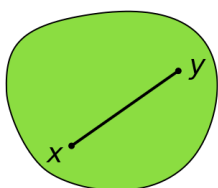


Abb. konvexe Mengen (Quelle 11)

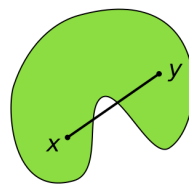


Abb: nicht konvexe Menge (Quelle 11)

Einfache Beispiele für konvexe Mengen sind unter anderem Einpunktmengen oder auch die leere Menge.

Um mit den konvexen Mengen arbeiten zu können, müssen wir zunächst auf einige Eigenschaften eingehen.

Proposition 3.13

Für konvexe Mengen gelten folgenden Aussagen:

1. Wenn zwei Mengen $A, B \subseteq \mathfrak{R}$ konvex sind, so ist auch die Summe $A + B$ wieder konvex.
2. Wenn eine Mengen $A \subseteq \mathfrak{R}$ konvex ist, so ist auch $\lambda * A$ wieder konvex für ein $\lambda \in \mathfrak{R}$.
3. Der Durchschnitt konvexer Mengen ist ebenfalls eine konvexe Menge.

Betrachten wir nun den Fall der konvexen Mengen auf Hadamard-Mannigfaltigkeiten. Hierbei wird der Begriff der Verbindungsline durch das geodätische Segment γ ersetzt, das nichts anderes darstellt, als die kürzeste Verbindungsline zwischen zwei beliebigen Punkten auf der Mannigfaltigkeit.

Definition 3.14

Betrachte eine Teilmenge C einer Hadamard-Mannigfaltigkeit M . Dann ist C genau dann konvex, wenn das geodätische Segment γ , welches zwei beliebige Punkte $x, x' \in C$ verbindet, ebenfalls in C enthalten ist. Dann gilt:

$$\gamma((1-t)a + tb) \in C \quad \forall t \in [0, 1].$$

Dabei wählen wir $x' = \gamma(a)$ und $x = \gamma(b)$.

Nachdem wir nun den Konvexitäts-Begriff auf Mengen eingeführt haben setzen wir für die folgenden Aussagen voraus, dass C eine abgeschlossene und konvexe Teilmenge von M ist. Das Ziel des folgenden Unterkapitels wird es sein, zu zeigen, dass es möglich ist einen Punkt $y_{x'}$ in der Teilmenge C zu finden, welcher den Abstand zu x' minimiert. Dieser Punkt wird auch Projektion von x' auf C genannt und wird mit $p_c(x')$ bezeichnet. Hierfür betrachten wir ein festes $x' \in C$ und definieren uns auf dieser Grundlagen die Menge $A_{x_0} = \{x \in M: d(x', x) \leq d(x', x_0)\}$. Betrachten wir den Schnitt von $A_{x_0} \cap C$ so stellen wir fest, dass dieser kompakt ist. Zudem wissen wir dass die Abbildung $x \rightarrow d(x', x)$ stetig ist.

Die in diesem Unterkapitel bearbeiteten Aussagen sind notwendig, um die Singularität der folgenden Aussage auf Basis der soeben eingeführten Punkte darzustellen:

Wenn die Karte $x \rightarrow d(x', x)$ stetig ist und die Menge $A_{x_0} \cap C$ kompakt, dann existiert ein Punkt $y_{x'} \in C$, sodass $d(x', y_{x'}) \leq d(x', x) \quad \forall x \in C$ gilt.

Dabei wird der Punkt $y_{x'}$ Projektion von x' auf C genannt und mit $p_c(x')$ bezeichnet, sodass für $x' \in M, x \in C$ die folgende Ungleichung gilt: $\langle \exp_{p_c(x')}^{-1} x', \exp_{p_c(x')}^{-1} x \rangle \leq 0$.

Diese Aussage werden wir indirekt über die Beweise der beiden folgenden Proposition beweisen, welche zusammengesetzt den gewünschten Beweis liefern.

Doch bevor wir die beiden Propositionen beweisen können müssen wir uns zunächst noch mit dem Thema Variationen auseinandersetzen und einige relevante Definitionen für den Beweis liefern:

Hierzu betrachten wir ein normalisiertes geodätisches Segment, welches gegeben ist durch die Abbildung $\gamma: [a, b] \rightarrow M$.

Eine differenzierbare Variation dieses geodätischen Segments ist dann definiert durch eine differenzierbare Karte $\gamma: [a, b] \times (-\varepsilon, \varepsilon) \rightarrow M$, welche die Bedingung $\alpha(t, 0) = \gamma(t)$ erfüllt.

Des Weiteren wird das Vektorfeld, welches entlang der Kurve von γ verläuft definiert durch $V(t) = \left(\frac{\partial \alpha}{\partial s}\right)(t, 0)$ und wird als Variations- Vektorfeld von α bezeichnet.

Zuletzt betrachten wir die erste Variationsformel von α welche gegeben ist durch :

$$L'(\gamma) := \frac{\partial}{\partial s} L(c_s) \Big|_{s=0} = \langle V, \gamma' \rangle \quad | \quad \text{mit } c_s = \alpha(t, s), s \in (-\varepsilon, \varepsilon).$$

Damit haben wir nun alle Fakten eingeführt, die wir nun im Beweis der folgenden Präposition benötigen .

Proposition 3.15

Sei $x' \in M$. Wenn $y_{x'} \in C$, sodass $d(x', y_{x'}) \leq d(x', x)$ gilt, dann gilt zusätzlich

$$\langle \exp_{y_{x'}}^{-1} x', \exp_{y_{x'}}^{-1} x \rangle \leq 0, \quad \forall x \in C.$$

Beweis:

Für den Fall, dass x' ebenso wie $y_{x'}$ in der Teilmenge $C \subset M$ liegt, ist unser Ergebnis trivial.

Aufgrund dessen betrachten wir den Fall $x' \notin C$ etwas genauer. Hierfür schaffen wir zunächst einige Voraussetzungen. Da wir wissen, dass das geodätische Segment genau dann minimal ist, wenn die Länge des geodätischen Segmentes die Riemannsche Distanz ist und der Punkt $y_{x'}$ jener sein soll, welcher die Länge minimiert, setzen wir zunächst

$l = d(x', y_{x'})$. Sei weiter die Geodäte beschrieben durch $\gamma: [0, l] \rightarrow M$, wobei dann $\gamma(0) = x'$

und $\gamma(l) = y_{x'} = \exp_{y_{x'}}^{-1} x'$ ist. Daraus folgt dann direkt $\gamma'(l) = \frac{-\exp_{y_{x'}}^{-1} x'}{\|\exp_{y_{x'}}^{-1} x'\|}$.

Um die eigentliche Aussage zu beweisen, führen wir einen Widerspruchsbeweis. Hierfür nehmen wir an, dass ein Punkt $\bar{x} \in C$ existiert , sodass $\langle \exp_{y_{x'}}^{-1} x', \exp_{y_{x'}}^{-1} \bar{x} \rangle > 0$ gilt.

Wenn wir im folgenden anhand der oben gegebenen Definition folgende Variation der Geodäte γ betrachten, welche gegeben ist durch $\alpha: [0, l] \times (-\varepsilon, \varepsilon) \rightarrow M$. Dabei setzen wir $\alpha(t, s) = \exp_{x'}^{-1} t(\exp_{x'}^{-1} \beta(s))$ mit $\beta: (-\varepsilon, \varepsilon + l) \rightarrow M$. Dabei wählen wir $\beta(0) = y_{x'}$ und $\beta(l) = \bar{x}$.

Des Weiteren wissen wir aus den gegebenen Definitione, dass das entsprechende Variation- Vektorfeld von α gegeben ist durch $V(t) = \left(\frac{\partial \alpha}{\partial s}\right)(t, 0)$. Dann wissen wir, dass das Variations Vektorfeld folgende Bedingungen erfüllt: $V(0) = 0$ und $V(l) = \beta'(0) = \exp_{x'}^{-1} \bar{x}$.

Wenn wir nun unter Verwendung der soeben gemachten Aussagen erste Variationsformel anwenden, erhalten wir folgende Aussage:

$$L'(Y) = \frac{\partial}{\partial s} L(c_s) \Big|_{s=0} = \left\langle \frac{-\exp_{y_{x'}}^{-1} x'}{\|-\exp_{y_{x'}}^{-1} x'\|}, -\exp_{y_{x'}}^{-1} \bar{x} \right\rangle < 0,$$

mit $c_s(\cdot) = \alpha(\cdot, s)$ und $s \in (-\varepsilon, \varepsilon)$.

Aus der letzten Ungleichung folgt, dass wenn ein $\delta > 0$ existiert, sodass $L(c_s) < L(Y)$ für beliebige $0 < s < \delta$, dann gilt $d(x', \beta(s)) \leq d(x', y_{x'})$. Da C konvex ist folgt, dass $\beta(s) \in C$ für beliebige $0 < s < \delta$.

Dann widerspricht die Ungleichung $d(x', \beta(s)) \leq d(x', y_{x'})$ der Definition von $y_{x'}$ und unser Beweis ist vollständig. □

Im nächsten Schritt werden wir nun beweisen, dass es genau einen Punkt in der Teilmenge C gibt, welcher den Abstand zu x' minimiert und dass dieser Punkt somit eindeutig ist. Hierzu formulieren wir zunächst die Aussage:

Proposition 3.16

Sei x' in M . Dann existiert genau ein $y_{x'}$ in C , welches den Abstand zu x' minimiert. Dann gilt $d(x', y_{x'}) \leq d(x', x)$ für alle $x \in C$.

Beweis:

Ebenso wie beim vorangegangenen Beweis ist der Fall für $x' \in C$ trivial, weshalb wir uns den Fall $x' \notin C$ anschauen. Um die Eindeutigkeit dieses Punktes zu zeigen, führen wir auch hier einen Widerspruchsbeweis, sodass wir zunächst annehmen, dass zwei Punkte $y_{x'}$ und $\overline{y_{x'}}$ in der Teilmenge C existieren, welche den Abstand zu x' minimieren, sodass gilt:

$$d(x', y_{x'}) = d(x', \overline{y_{x'}}) \leq d(x', x), \forall x \in M.$$

Wenn wir nun das geodätische Dreieck $\Delta(x', y_{x'}, \overline{y_{x'}})$ betrachten und deren Winkel mit $\theta, \bar{\theta}, \theta'$ bezeichnen, dann erhalten wir mit der bereits im einführenden Kapitel gegebenen Definition für

$$\text{die Winkel } \theta = \angle(\exp_{y_{x'}}^{-1} x', \exp_{y_{x'}}^{-1} \overline{y_{x'}}), \bar{\theta} = \angle(\exp_{\overline{y_{x'}}}^{-1} x', \exp_{\overline{y_{x'}}}^{-1} y_{x'}),$$

$$\theta' = \angle(\exp_{x'}^{-1} y_{x'}, \exp_{x'}^{-1} \overline{y_{x'}}). \text{ Da } y_{x'} \neq \overline{y_{x'}} \text{ wissen wir, dass } \theta \geq \frac{\pi}{2}, \bar{\theta} \geq \frac{\pi}{2}.$$

Da wir jedoch auch angenommen haben, dass die Riemannsche Distanz von x' zu $y_{x'}$ gleich der Distanz von x' zu $\overline{y_{x'}}$ ist, d.h., dass $d(x', y_{x'}) = d(x', \overline{y_{x'}})$ gilt und wir wissen das C konvex ist, folgt daraus direkt dass $\theta' > 0$ sein muss. Dies widerspricht jedoch die Aussage aus Theorem 3.11, welche aussagt, dass $\theta + \theta' + \overline{\theta} > \pi$. Daraus folgt der Beweis. \square

Nachdem wir nun die beiden Propositionen bewiesen haben, können wir mit dem Wissen aus der Anmerkung zu Beginn recht einfach erkennen, dass das folgende Korollar direkt aus den vorangegangenen Aussagen folgt.

Korollar 3.17

Sei x' in M . Dann existiert eine eindeutige Projektion $p_C(x')$ und es gilt folgende Ungleichung für alle $x \in C$:

$$\langle \exp_{p_C(x')}^{-1} x', \exp_{p_C(x')}^{-1} x \rangle \leq 0.$$

Nachdem wir uns nun mit dem Thema der Projektion beschäftigt haben, kommen wir nun zum letzten Thema dieses Teilabschnitts. Hier werden wir uns mit dem Support beschäftigen. Die Ergebnisse dieses Abschnittes werden im weiteren Verlauf eine tragende Rolle spielen. Unter anderem benötigen wir sie für die Beweisführung eines sehr relevanten Theorems, aus welchem wir direkt die Wohldefiniertheit des Proximal-Punkt Algorithmus erhalten.

Wir werden zunächst den Support definieren und anschließend zwei wichtige Aussagen hierzu angeben.

Definition 3.18

Angenommen der Rand ∂C der abgeschlossenen konvexen Menge C ist nichtleer und sei $x' \in \partial C$ und $s \in T_{x'} M$, wobei $s \neq 0$.

Dann definieren wir einen Unterraum von $T_{x'} M$ durch $S_{s,x'} := \{v \in T_{x'} M : \langle s, v \rangle = 0\}$.

Dieser Unterraum heißt Support von C im Punkt x' , genau dann wenn $\langle s, \exp_{x'}^{-1} x \rangle \leq 0$ ist für alle $x \in C$.

Lemma 3.19

a) Sei $\bar{x} \in M$, sodass jedoch $\bar{x} \notin C$. Genau dann existiert ein $s \in M$, sodass

$$\sup\{\langle -s, \exp_x^{-1} x \rangle : x \in C\} < 0.$$

b) Sei $x' \in \partial C$. Dann existiert ein Support-Unterraum zu C im Punkt x' .

Beweise zu diesen Aussagen finden wir in [1, S.263. f.]

3.2.2 Einführung in das Gebiet der konvexen Funktionen

Dieses Unterkapitel beinhaltet neben der Definition von konvexen Funktionen ebenso einige wichtige Eigenschaften. Hierbei werden wir die ein oder andere Eigenschaft beweisen, wenn diese im Kontext eine relevante Rolle spielen wird. Abschließend werden wir uns in diesem Unterkapitel mit der Definition des Subgradienten beschäftigen, da dieser seine Anwendung im folgenden Unterkapitel findet.

Wir beginnen mit der Definition einer konvexen Funktion. Hierbei wählen wir den geometrischen Weg, diese über ihren Epigraphen zu definieren.

Definition 3.20

Der Epigraph einer Funktion $f: K \rightarrow \mathfrak{R}$ ist gegeben durch die Menge

$$Epi(f) = \{(x, r) \in \mathfrak{R}^n \times \mathfrak{R} : x \in K, r \geq f(x)\}.$$

Betrachtet man eine Funktion $f: M \rightarrow \mathfrak{R}$, wobei M eine riemannsche Mannigfaltigkeit ist, dann sieht der Epigraph wie folgt aus:

$$Epi(f) = \{(x, r) \in M \times \mathfrak{R} / f(x) \geq r\}.$$

Dieser stellt einen konvexen Unterraum in $M \times \mathfrak{R}$ dar.

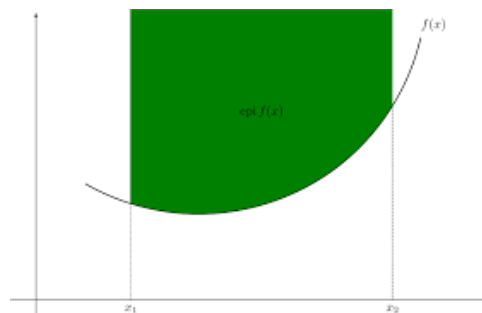


Abb: Epigraph von $f(x)$ - (Quelle 11)

Anhand der Definition des Epigraphen kann man nun Aussagen über die Konvexität von Funktionen machen.

Satz 3.21

Sei M eine riemannsche Mannigfaltigkeit. Dann heißt eine Funktion $f: M \rightarrow \mathfrak{R}$ konvex, wenn ihr Epigraph konvex ist.

Beweis:

" \Rightarrow " Seien $x, y \in \text{Epi}(f)$ mit $x = (x_1, x_2)$, $y = (y_1, y_2)$ und $\lambda \in [0, 1]$, wobei $x_1, y_1 \in M$ und $x_2, y_2 \in \mathfrak{R}$. Sei $z = (z_1, z_2) := \lambda x + (1 - \lambda)y = (\lambda x_1 + (1 - \lambda)y_1, (\lambda x_2 + (1 - \lambda)y_2))$.

Dann gilt $z_2 = \lambda x_2 + (1 - \lambda)y_2 \geq \lambda f(x_1) + (1 - \lambda)f(y_2) \geq f(\lambda x_1 + (1 - \lambda)y_1) = z_1$.

Daraus folgt: $z \in \text{Epi}(f)$.

" \Leftarrow " Sei $\text{Epi}(f)$ konvex gegeben mit $(f(x), x), (f(y), y) \in \text{Epi}(f)$ und $\lambda \in [0, 1]$. Dann gilt $f(\lambda x + (1 - \lambda)y) > \lambda f(x) + (1 - \lambda)f(y)$ und somit ist f nicht konvex. Daraus folgt: $\lambda(x, f(x)) + (1 - \lambda)(y, f(y)) = (\lambda x + (1 - \lambda)y, \lambda f(x) + (1 - \lambda)f(y)) \notin \text{Epi}(f)$. Somit wäre auch $\text{Epi}(f)$ nicht konvex, woraus ein Widerspruch folgt. \square

Bemerkung 3.22

1. Eine konvexe Funktion hat als Definitionsmenge immer eine konvexe Menge. Ist die Definitionsmenge nicht konvex, so ist auch die Funktion nicht konvex.
2. Wenn f eine konvexe Funktion ist, dann ist jedes lokale Minimum auch ein globales Minimum
3. Wenn f konvex ist, dann besitzt die Funktion höchstens ein lokales Minimum.

Beweis zu 3. :

Gegeben zwei lokale Minima x, y von f mit $x \neq y$. Da aus Punkt zwei der Bemerkung bereits bekannt ist, dass alle lokalen Minima auch globale Minima sind, sind auch x, y mit $f(x) = f(y)$ globale Minima.

Mit der analytischen Definition einer konvexen Funktion folgt dann: $f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y) = f(x) = f(y)$. Hieraus folgt direkt ein Widerspruch und somit direkt die Aussage. \square

Weitere Begriffe, welche im Zusammenhang mit diesem Thema nicht unterschlagen werden dürfen und welche im Folgenden immer wieder auftauchen und genutzt werden, sind der

Begriff des Subgradienten und des Subdifferentials. Daher wollen wir diese zunächst definieren:

Definition 3.23

Sei eine konvexe Funktion $f: M \rightarrow \mathfrak{R}$ gegeben mit M ist Hadamard-Mannigfaltigkeit und wähle $x' \in M$. Dann ist $s \in T_{x'}M$ ein Subgradient der Funktion f im Punkt $x' \in M$, falls die Ungleichung $\forall x \in M$ gilt:

$$f(x) \geq f(x') + \langle s, \exp_{x'}^{-1} x \rangle.$$

Das Subdifferential ist bekannterweise eine Verallgemeinerung des Gradienten auf dem Gebiet der nicht differenzierbaren konvexen Funktionen und wird in dieser Ausarbeitung mit $\partial f(x')$ bezeichnet.

Damit die soeben eingeführten Begrifflichkeiten auch direkt ihre Anwendung finden, kommen wir als nächstes zu einem Theorem, welches aussagt, dass das Subdifferential $\partial f(x')$ von einer konvexen Funktion f auf einer Hadamard-Mannigfaltigkeit nicht leer ist.

Das Theorem findet seine Anwendung in der Beweisführung zur Wohldefiniertheit der Moreau-Yosida Regularisierung.

Aufgrund der Relevanz dieses Theorems werden wir hier erweiternd auch den Beweis zeigen.

Theorem 3.24

Sei M eine Hadamard Mannigfaltigkeit und sei $f: M \rightarrow \mathfrak{R}$ konvex. Dann ist für beliebige $x' \in M, s \in T_{x'}M$, sodass $f(x) \geq f(x') + \langle s, \exp_{x'}^{-1} x \rangle, \forall x \in M$.

Beweis:

Aus den vorangegangenen Sätzen und Definitionen zum Epigraphen wissen wir, dass es sich bei $Epi(f)$ um eine geschlossene, konvexe Teilmenge von $M \times \mathfrak{R}$ handelt, wobei f stetig ist.

Sei die Grenze von $Epi(f)$ gegeben durch $\partial(epi(f)) = \{(s, f(x)): x \in M\}$ und die exponentielle Karte auf $M \times \mathfrak{R}$ im Punkt $(x', f(x'))$ definiert durch

$$\exp_{(x', f(x'))}^{-1}(x, r) := (\exp_{x'}^{-1} x, r - f(x')).$$

Da die Grenze von $Epi(f)$ nicht leer ist, existiert ein Support-Unterraum, welcher nicht trivial ist. Dieser Support-Unterraum von $Epi(f)$ im Punkt $(x', f(x'))$ wird wie folgt bezeichnet:

$S_{(s,\alpha), (x',f(x'))}$. Daraus folgt, dass $\forall(x, t) \in \text{Epi}(f)$ gilt:

$$\langle s, \exp_{x'}^{-1} x \rangle + \alpha(r - f(x')) = \langle (s, \alpha), (\exp_{x'}^{-1} x, r - f(x')) \rangle \leq 0.$$

Bezeichnen wir im folgenden $\bar{x} = \exp_{x'}^{-1} s$ und $\bar{r} = f(\bar{x})$ und setzen $x = \bar{x}$ und $r = \bar{r}$.

Dann folgt aus der oben genannten Gleichung:

$$\begin{aligned} & \langle s, \exp_{x'}^{-1} \bar{x} \rangle + \alpha(\bar{r} - f(x')) \leq 0 \\ \Leftrightarrow & \langle (s, \alpha), (\exp_{x'}^{-1} \bar{x}, \bar{r} - f(x')) \rangle \leq 0 \\ \Leftrightarrow & \langle (s, \alpha), (\exp_{x'}^{-1} \exp_{x'} s, \bar{r} - f(x')) \rangle \leq 0 \\ \Leftrightarrow & \|s\|^2 + \alpha(\bar{r} - f(x')) \leq 0 \\ \Leftrightarrow & \|s\|^2 \leq \alpha(f(x') - \bar{r}). \end{aligned}$$

Da wir ja bereits wissen, dass der definierte Support-Unterraum nicht trivial also auch nicht leer ist, folgt direkt, dass $\alpha \neq 0$ sein muss. Darum nehmen wir ohne Einschränkung der Allgemeinheit an, dass $\alpha = (-1)$. Wenn wir jetzt noch die Annahme treffen, dass $r = f(x)$ und dies dann auf die Ursprungsgleichung anwenden, folgt

$$\begin{aligned} & \langle s, \exp_{x'}^{-1} x \rangle + \alpha(r - f(x')) \leq 0 \\ \Leftrightarrow & \langle s, \exp_{x'}^{-1} x \rangle + (-1)(f(x) - f(x')) \leq 0 \\ \Leftrightarrow & f(x) \geq f(x') + \langle s, \exp_{x'}^{-1} x \rangle \quad \forall x \in M \end{aligned}$$

und somit direkt die gewünschte Aussage. □

3.3 Regularisierung

Der Prozess der Regularisierung spielt in der Mathematik eine tragende Rolle und wird in den verschiedensten Bereichen verwendet. Sie wird unter anderem bei schlecht gestellten Optimierungsproblemen eingesetzt, um Prozesse zu vereinfachen oder zu stabilisieren. Dazu werden spezielle Datenterme erzeugt, anhand dessen man abschätzen kann, wie sehr man sich an die vorhandenen Daten hält oder ob doch eine Verallgemeinerung notwendig bzw. besser ist.

In unserem Fall heißt diese Regularisierung, welche zur Stabilisierung des Optimierungsproblems erzeugt wurde, "Moreau-Yosida-Regularisierung".

Sie wird uns helfen, bessere Eigenschaften für unseren Proximal-Punkt-Algorithmus zu erhalten.

Wir wollen dieses sogenannte Skalierungsparameter in diesem Unterkapitel einführen und im Anschluss direkt die Wohldefiniertheit dieser Regularisierung zeigen.

3.3.1 Die Moreau-Yosida Regularisierung

Bevor wir in diesem Unterkapitel zur Moreau-Yosida-Regularisierung kommen, müssen wir zunächst noch einen Blick auf die Definition der inversen Karte werfen, sowie auf den Gradienten dieser. Die Definition der inversen Karte ist Bestandteil der Moreau-Yosida-Regularisierung, weshalb wir diese vorab einführen werden:

Definition 3.25

Sei M eine Hadamard-Mannigfaltigkeit. Die inverse Karte $\exp_{x'}^{-1}: M \rightarrow T_{x'}M$ ist für beliebige $x' \in M$ unendlich oft stetig differenzierbar. Wenn weiter $d(x', x) = \|\exp_x^{-1}x'\|$ gilt, dann wird die Karte $p_{x'}: M \rightarrow \mathfrak{R}$ definiert durch $p_{x'}(x) = \frac{1}{2}d^2(x, x')$, welche ebenfalls C^∞ ist.

Die folgende Proposition gilt als Ergänzung zu der soeben gegebenen Definition der inversen Karte und macht dabei eine Aussage über ihren Gradienten:

Proposition 3.26

Betrachten wir wiederum die Hadamard-Mannigfaltigkeit M . Wenn $x' \in M$, dann ist $p_{x'}: M \rightarrow \mathfrak{R}$ strikt konvex und der Gradient der inversen Karte $p_{x'}(x) = \frac{1}{2}d^2(x, x')$ im Punkt x ist definiert durch $p_{x'} = -\exp_x^{-1}x'$.

Nachdem wir nun zwei für die Definition der Moreau-Yosida-Regularisierung wichtige Aussagen eingeführt haben, wollen wir nun diese definieren:

Definition 3.27

Sei M eine Hadamard-Mannigfaltigkeit und $f: M \rightarrow \mathfrak{R}$ eine konvexe Funktion. Die Moreau-Yosida-Regularisierung $f_\lambda: M \rightarrow \mathfrak{R}$ mit $\lambda > 0$ ist dann definiert durch:

$$f_\lambda(x) = \inf_{y \in M} \left\{ f(y) + \lambda p_x(y) \right\}, \text{ wobei } p_x \text{ gegeben ist durch } p_x(x) = \frac{1}{2}d^2(x, x').$$

3.3.2 Wohldefiniertheit der Moreau-Yosida Regularisierung

Um die Wohldefiniertheit der oben aufgeführten Moreau-Yosida-Regularisierung zu zeigen, müssen wir zunächst den Begriff der Koerzivität einführen. Dies ist nötig, um bessere Aussagen zu der Eindeutigkeit und Existenz eines Minimums treffen zu können. Da stetige Funktionen im Allgemeinen auf nicht kompakten Mengen ihr Minimum nicht annehmen, wollen wir uns einen Spezialfall zu Nutze machen. Wenn eine stetige Funktion zusätzlich noch die Eigenschaften der Konvexität und der Koerzivität erfüllt, dann nimmt diese Funktion ihr Minimum immer an.

Daher kommen wir nun zur Definition:

Definition 3.28

Eine Funktion $h: M \rightarrow \mathfrak{R}$ ist 1-koerziv in $x \in M$, wenn $\lim_{d(x,y) \rightarrow +\infty} \frac{h(y)}{d(x,y)} = +\infty$

Falls $h: M \rightarrow \mathfrak{R}$ eine 1-coercive Funktion in $x \in M$, dann ist einfach zu zeigen, dass die minimierende Menge von h nicht leer ist.

Bemerkung 3.29

Koerzive Funktionen haben mindestens ein Minimum.

Lemma 3.30

Wenn die Funktion $f: M \rightarrow \mathfrak{R}$ konvex ist und $\lambda > 0$, dann ist die Funktion $(f + \lambda \rho_x): M \rightarrow \mathfrak{R}$ 1-koerziv in $x \in M$.

Beweis:

Da die Funktion $f: M \rightarrow \mathfrak{R}$ konvex ist, können wir auf eine bereits bewiesene Aussage zurückgreifen. Diese besagt, dass wenn M eine Hadamard-Mannigfaltigkeit ist und f konvex ist, dass dann für beliebige $x' \in M$ ein Subgradient $s \in T_x M$ von f in x' existiert, sodass gilt:

$f(x) \geq f(x') + \langle s, \exp(x')^{-1} x \rangle \quad \forall x \in M$. Wenden wir diese Aussage nun auf unsere Funktion $(f + \lambda \rho_x): M \rightarrow \mathfrak{R}$ an, so erhalten wir ein $s \in T_x M$, sodass:

$$\begin{aligned} \frac{(f + \lambda \rho_x)(y)}{d(x,y)} &\geq \frac{f(x)}{d(x,y)} + \langle s, \frac{\exp_x^{-1} y}{d(x,y)} \rangle + \lambda \frac{\rho_x(y)}{d(x,y)} \\ &\geq \frac{f(x)}{d(x,y)} + \langle s, \frac{\exp_x^{-1} y}{d(x,y)} \rangle + \lambda \frac{\frac{1}{2} d^2(x,y)}{d(x,y)} \end{aligned}$$

$$\geq \frac{f(x)}{d(x,y)} + \langle s, \frac{\exp_x^{-1} y}{d(x,y)} \rangle + \frac{\lambda}{2} d(x,y)$$

Ferner haben wir bereits $d(x,y) = \left\| \exp_x^{-1} y \right\|$ definiert und $\lambda > 0$ nach Aussage.

Daraus folgt dann wiederum mit Hilfe der Definition der Koerzivitat:

$$\lim_{d(x,y) \rightarrow +\infty} \frac{(f + \lambda p_x)(y)}{d(x,y)} = +\infty. \quad \square$$

Lemma 3.31

Wenn $f: M \rightarrow \mathfrak{R}$ konvex ist, dann existiert fur beliebige $x \in M$ und $\lambda > 0$ ein eindeutiger Punkt, welcher mit $p_\lambda(x)$ bezeichnet wird, sodass

$$f(p_\lambda(x)) + \lambda p_x(p_\lambda(x)) = f_\lambda(x)$$

charakterisiert durch

$$\lambda(\exp_{p_\lambda(x)}^{-1} x) \in \partial f(p_\lambda(x)).$$

Beweis:

Aus vorangegangenen, bereits bewiesenen Aussagen, wissen wir, dass p_x streng konvex ist. Da wir ebenfalls wissen, dass jede Positivkombination von konvexen Funktionen, also jede Summe bzw. jede Multiplikation einer konvexen Funktion mit einer reellen Zahl, ebenfalls konvex ist, konnen wir hieraus direkt die Konvexitat von $f + \lambda p_x$ folgern. Somit ist $f + \lambda p_x$ streng konvex.

In einer weiteren Bemerkung, haben wir bereits gezeigt, dass jede streng konvexe Funktion hochstens ein lokales Minimum hat, welches auch gleichzeitig das globale Minimum ist. Somit hat $f + \lambda p_x$ hochstens ein Minimum aufgrund der Konvexitat.

Ferner wurde bewiesen, dass wenn eine Funktion $f: M \rightarrow \mathfrak{R}$ konvex ist und $\lambda > 0$ ist, dass $f + \lambda p_x$ 1-koerziv ist. Bei koerziven Funktionen ist bekannt, dass diese mindestens ein Minimum besitzen, welches sie nur annehmen, wenn die Bedingungen der Stetigkeit und der Konvexitat erfullt werden. Da dies hier der Fall ist, folgt direkt, dass $f + \lambda p_x$ genau ein eindeutiges Minimum annimmt.

Durch die Eindeutigkeit des Minimums folgt direkt aus der Moreau-Yosida-Regulierung $f(p_x(x)) + \lambda p_x(p_x) = f_\lambda(x)$ für die Existenz eines Eindeutigen Punktes von $p_\lambda(x)$. Für $p_\lambda(x)$ Minimum von $f + \lambda p_x$ folgt direkt, dass $0 \in \partial(f + \lambda p_x)(p_\lambda(x))$ und damit auch die Charakterisierung für $p_\lambda(x)$, welche dann gegeben ist durch:

$$\partial(f + \lambda p_x)(y) = \partial f(y) + \{-\lambda \exp_x^{-1} x\}. \quad \square$$

Kapitel 4

Der Proximal-Punkt Algorithmus

Der Inhalt dieses Kapitels orientiert sich an [1, S.267 f.] ,[3, S.1 f.] , [4, S.1 f.], [9, S.1 f.].

4.1 Einführung des Proximal-Punkt Algorithmus

In diesem Kapitel gehen wir auf das eigentliche Hauptthema dieser Ausarbeitung ein, dem Proximal-Punkt-Algorithmus. Bereits in Kapitel zwei, in dem wir das Optimierungsproblem eingeführt haben, haben wir die Iteration dieses Verfahrens dargestellt. Um nun in diesem Kapitel näher auf dieses Thema einzugehen, haben wir uns zunächst mit der Riemannschen Geometrie und der konvexen Analysis beschäftigt, anhand dessen wir uns die grundlegenden Themen angeschaut haben, um uns nun mit der Wohldefiniertheit und der Konvergenz dieses Verfahrens zu beschäftigen.

Eine wichtige Rolle spielt hierbei die in Kapitel drei eingeführte Moreau-Yosida Regularisierung. Diese stellt die Grundlage für den Proximal Punkt-Algorithmus, da sich die Iteration des Proximal-Punkt Algorithmus direkt aus der Moreau-Yosida-Regularisierung und der Definition der Folge $\{x_k\}$ zusammensetzt.

Wir betrachten den Proximal-Punkt Algorithmus auf riemannschen Mannigfaltigkeiten als eine natürliche Verallgemeinerung des durch Martinet eingeführten Proximal- Algorithmus auf dem \mathfrak{R}^n .

Betrachten wir zunächst den Fall, dass unsere Hadamard-Mannigfaltigkeit $M = \mathfrak{R}^n$ ist. Dann wäre das Verfahren durch die folgende Iteration gegeben:

$$x_{k+1} = \operatorname{argmin}_{y \in \mathfrak{R}^n} \left\{ f(y) + \lambda_k \|x_k - y\|^2 \right\}$$

wobei $p_{x_k}(x) = \|x_k - x\|$.

Wenn wir den Proximal-Punkt Algorithmus nun ohne die Einschränkung, dass $M = \mathfrak{R}^n$ betrachten kommen wir zur folgenden Definition:

Definition 4.1

Sei M eine Hadamard-Mannigfaltigkeit zusammen mit einer konvexen Funktion $f: M \rightarrow \mathfrak{R}$. Ferner sei $p_{x_k}: M \rightarrow \mathfrak{R}$ eine Karte, welche definiert ist durch $p_{x_k}(x) = \frac{1}{2}d^2(x, x_k)$. Dann wird durch den Proximal-Punkt Algorithmus eine Folge $\{x_k\} \subset M$ erzeugt, welche für einen Startpunkt $p_0 \in M$ und einer Folge positiver Zahlen $\{\lambda_k\}$ gegeben wird durch folgende Iteration:

$$\begin{aligned}x_{k+1} &= p_{\lambda_k}(x_k) \\ &= \operatorname{argmin}_{y \in M} \{f(y) + \lambda_k p_{x_k}(y)\}.\end{aligned}$$

Der Proximal-Punkt Algorithmus ist ein Algorithmus um Minima einer konvexen Funktion f zu berechnen, indem er die Nullstellen eines Teildifferentials der konvexen Funktion f bestimmt. Dabei macht sich der Algorithmus nutzen von der Tatsache, dass das Subdifferential ein monotoner Operator ist. Die Besonderheit unseres Algorithmus ist, dass dieser eingeschränkt auch nicht konvexe Problemstellung auf dem \mathfrak{R}^n löst. Dies geschieht unter der Bedingung, dass die Menge, auf welche eingeschränkt wird, eine Riemannsche Mannigfaltigkeit ist, genauer eine Hadamard-Mannigfaltigkeit. Die Einschränkung auf eine Hadamard-Mannigfaltigkeit muss vorgenommen werden, da wir für den Beweis der Konvergenz unbedingt eine nicht positive Schnittkrümmung des vorliegenden Raumes haben müssen.

Bevor wir jedoch zu Konvergenz kommen, kommen wir zur Wohldefiniertheit des Algorithmus.

4.2 Wohldefiniertheit des Proximal-Punkt Algorithmus

Wie bereits in den vorherigen einleitenden Worten erwähnt, spielt die Moreau-Yosida-Regularisierung für den Proximal-Punkt-Algorithmus eine elementare Rolle. In diesem Abschnitt wollen wir uns nun mit der Wohldefiniertheit des Proximal-Punkt-Algorithmus beschäftigen. Die Beweisführung läuft in diesem Falle sehr analog zu dem bereits gezeigten Beweis der Wohldefiniertheit der Moreau-Yosida-Regularisierung.

Theorem 4.2

Sei $f: M \rightarrow \mathbb{R}$ eine konvexe Funktion, wobei M eine Hadamard Mannigfaltigkeit ist. Die Folge $\{x_k\}$, welche durch den Algorithmus

$$x_{k+1} = p_{\lambda_k}(x_k) = \operatorname{argmin}_{y \in M} \{f(y) + \lambda_k p_{x_k}(y)\}$$

erzeugt wurde, ist wohldefiniert und charakterisiert durch

$$\lambda_k (\exp_{x_{k+1}}^{-1} x_k) \in \partial f(x_{k+1}).$$

Beweis:

Analog zum Beweis der Wohldefiniertheit der Moreau-Yosida-Regularisierung wird der Beweis der Eindeutigkeit des Proximal-Punkt-Algorithmus geführt. Ein Unterschied besteht darin, dass der Beweis zuvor anhand eines Punktes geführt wurde. Jetzt steht an dessen Stelle eine Folge $\{x_k\} \subset M$ anhand derer die Wohldefiniertheit der Iteration gezeigt werden soll.

Setze $p_{\lambda_k}(x_k) = x_{k+1}$ und sei $p_{x_k}(x) = \frac{1}{2}d^2(x_k, x)$ und $\{\lambda_k\}$ eine Folge positiver Zahlen. Wir wissen auch hier, dass p_x streng konvex ist und daher auch $f + \lambda p_x$ streng konvex. Analog folgt hier die Koerzivität. Daher hat $f + \lambda p_x$ ein eindeutiges Minimum und $p_{\lambda_k}(x_k) = x_{k+1} = \operatorname{argmin}_{y \in M} \{f(y) + \lambda_k p_{x_k}(y)\}$ folgt direkt aus der Moreau-Yosida-Regularisierung. Bezeichnet $p_{\lambda_k}(x_k)$ die minimierende Folge von $f + \lambda p_x$ folgt daraus direkt $0 \in \partial (f + \lambda p_x)(y)$. \square

4.3. Konvergenz

Nachdem wir im Abschnitt zuvor die Wohldefiniertheit des Proximal-Punkt-Algorithmus gezeigt haben, kommen wir in diesem letzten Abschnitt zur Betrachtung der Konvergenz der durch die Iteration erzeugten Folge.

Um die Konvergenz für dieses Iterationsverfahren zeigen zu können, müssen wir zunächst einige Definitionen, Sätze und deren Beweise betrachten. Hierzu sei die durch die Iteration erzeugte Folge $\{x_k\} \subset M$ zu betrachten.

Eines der Kriterien, welches für die Konvergenz erfüllt sein muss, ist, dass die Folge begrenzt sein muss. Wenn wir einen vollständigen metrischen Raum (M, d) betrachten,

können wir beweisen, dass die Folge $\{x_k\}$ beschränkt ist, wenn diese Fejer-konvergent ist. Doch zuvor müssen wir hierfür definieren was Fejer-Konvergenz überhaupt für uns bedeutet:

Definition 4.3

Sei (M, d) ein vollständiger metrischer Raum. Die Folge $\{x_k\} \subset M$ ist genau dann Fejer-konvergent zu einer nichtleeren Menge $U \subset M$, wenn gilt:

$$d(x_{k+1}, y) \leq d(x_k, y) \quad \forall y \in U, k \geq 0.$$

Mit Hilfe dieser Definition können wir uns nun auf die bereits oben angesprochene Beschränktheit der Folge $\{x_k\}$ konzentrieren.

Lemma 4.4

Sei (M, d) ein vollständiger metrischer Raum. Wenn $\{x_k\} \subset M$ Fejer-konvergent zu einer nichtleeren Teilmenge $U \subseteq M$ ist, dann ist die Folge $\{x_k\}$ beschränkt.

Ferner, falls ein Cluster Punkt x von $\{x_k\}$ zu U gehört, dann ist $\lim_{k \rightarrow \infty} x_k = x$.

Beweis:

Wir wählen $u \in U$. Wenn wir nun die soeben gelieferte Definition der Fejer-Konvergenz betrachten, genauer genommen die Ungleichung $d(x_{k+1}, y) \leq d(x_k, y)$, welche die Fejer-Konvergenz impliziert folgt für unseren Fall die folgende Ungleichung: $d(x_k, u) \leq d(x_0, u) \forall k$ und alle $u \in U$, d.h. dass die Riemannsche Distanz zwischen unserem gewählten $u \in U$ und x_k geringer oder gleich der Distanz zwischen x_0 und u ist. Diese Ungleichung impliziert direkt, dass die Folge $\{x_k\}$ beschränkt ist.

Da (M, d) ein vollständiger metrischer Raum ist und die Folge $\{x_k\}$ beschränkt, existiert eine konvergente Teilfolge. Im Folgenden betrachten wir also eine Teilfolge $\{x_{k_j}\}$ von $\{x_k\}$, sodass $\lim_{k \rightarrow +\infty} x_{k_j} = x$.

Da $x \in U$ haben wir durch $\{d(x_k, x)\}$ eine abnehmende Folge von positiven Zahlen gegeben. Diese besitzt wiederum eine gegen null konvergierende Teilfolge $\{d(x_{k_j}, x)\}$.

Dies impliziert wiederum, dass die gesamte Folge gegen null konvergiert, d.h. $0 = \lim_{k \rightarrow \infty} d(x_k, x) \Rightarrow x = \lim_{k \rightarrow \infty} x_k$.

Daraus folgt die gewünschte Aussage □

Ein ebenso für den Beweis der Konvergenz relevante Ungleichung werden wir im folgenden Lemma einführen und beweisen.

Lemma 4.5

Sei M eine Hadamard Mannigfaltigkeit und sei $f: M \rightarrow \mathbb{R}$ eine konvexe Funktion. Falls die Folge $\{x_k\}$ durch die Iteration des Proximal-Punkt Algorithmus

$$x_{k+1} = p_{\lambda_k}(x_k) = \operatorname{argmin}_{y \in M} \{f(y) + \lambda_k p_{x_k}(y)\}$$

erzeugt wird, dann gilt die folgende Ungleichung

$$d^2(x_{k+1}, x_k) \leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + \frac{2}{\lambda_k} (f(x) - f(x_{k+1})) \quad \forall x \in M$$

Beweis:

Wir betrachten das geodätische Dreieck $\Delta(x_k, x_{k+1}, x)$ und wählen $x \in M$. Betrachten wir nun das Theorem 3.11, können wir folgende Ungleichung aufstellen: Betrachten wir nun die in Theorem 3.11 eingeführte Aussage. Wir wissen, dass die geodätischen Segmente die Punkte unseres geodätischen Dreiecks $\Delta(x_k, x_{k+1}, x)$ verbinden. Diese werden wie folgt bezeichnet: $\gamma_{i+1}: [0, l_{i+1}] \rightarrow M$. Dann definieren wir $l_{i+1} := L(\gamma_{i+1})$. Minimieren wir das Längenfunktional L über allen Kurven so erhalten wir die Riemannsche Distanz und somit folgt aus der Ungleichung aus Theorem 3.11 folgendes:

$$d^2(x_k, x_{k+1}) + d^2(x_{k+1}, x) - 2d(x_k, x_{k+1})d(x_{k+1}, x)\cos(\theta) \leq d^2(x_k, x).$$

Nun betrachten wir die Definition des Winkels in Theorem 3.11, welcher sich wie folgt darstellen lässt: $\theta_{i+1} = \angle(\gamma'_{i+1}(0), -\gamma'_i(l_i))$. Für unseren Beweis wählen wir den Winkel $\theta = \angle(\exp_{x_{k+1}}^{-1} x_k, \exp_{x_{k+1}}^{-1} x)$. Betrachten wir nun das Skalarprodukt dieses Winkels, erhalten

wir die folgende Aussage: $\langle \exp_{x_{k+1}}^{-1} x_k, \exp_{x_{k+1}}^{-1} x \rangle = d(x_k, x_{k+1})d(x_{k+1}, x)\cos(\theta)$.

Ersetzen wir nun $d(x_k, x_{k+1})d(x_{k+1}, x)\cos(\theta)$ in unserer Ungleichung erhalten wir die folgende Aussage:

$$d^2(x_k, x_{k+1}) + d^2(x_{k+1}, x) - 2 \langle \exp_{x_{k+1}}^{-1} x_k, \exp_{x_{k+1}}^{-1} x \rangle \leq d^2(x_k, x).$$

Aus Theorem 4.2 wissen wir, dass $\lambda_k(\exp_{x_{k+1}}^{-1} x_k) \in \partial f(x_{k+1})$ und in Kombination mit unserer Ungleichung folgt dann folgende Aussage:

$$\begin{aligned}
d^2(x_{k+1}, x) &\leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + 2 \langle \exp_{x_{k+1}}^{-1} x_k, \exp_{x_{k+1}}^{-1} x \rangle \\
&\leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + \frac{2}{\lambda_k} (f(x) - f(x_{k+1}))
\end{aligned}$$

Dies zeigt die gewünschte Aussage. □

Die soeben gezeigten Aussagen spielen eine relevante Rolle für das nächste Theorem, in welchem wir zeigen wollen, dass die durch $\{x_k\}$ generierte Folge konvergiert.

Theorem 4.6

Sei $\{x_k\}$ eine durch (24) erzeugte Folge. Wenn die Folge $\{\lambda_k\}$ so ist, dass $\sum_{k=0}^{\infty} \left(\frac{1}{\lambda_k}\right) = +\infty$,

dann ist $\lim_{k \rightarrow \infty} f(x_k) = f^*$, wobei $f^* = \inf_{x \in M} f(x)$. Wenn zusätzlich die minimierende Menge

U^* nichtleer ist, dann $\lim_{k \rightarrow \infty} x_k = x_*$ und $x_* \in U^*$

Beweis:

Die Grundlage für diesen Beweis bilden die beiden bereits zuvor bewiesenen Aussagen. Wir werden bei diesem Beweis in zwei Schritten vorgehen, wobei wir zunächst zeigen werden, dass die durch den Proximal-Punkt Algorithmus generierte Folge $\{x_k\}$ Fejer-konvergent ist. Hierbei werden wir vor allem Lemma 4.5 zur Hilfe nehmen. Im zweiten Schritt werden wir dann mit Hilfe von Lemma 4.5 und weiteren bekannten Aussagen zeigen, dass der Clusterpunkt $x_* \in U^*$ ist.

Beginnen wir mit der Fejer-Konvergenz der Folge $\{x_k\}$. Sei dazu $x_k \notin U^*$. Nun betrachten wir die in Lemma 4.5 dargestellte Ungleichung und ersetzen in dieser alle x durch x_k und erhalten folgendes:

$$\begin{aligned}
d^2(x_{k+1}, x_k) &\leq d^2(x_k, x_k) - d^2(x_k, x_{k+1}) + \frac{2}{\lambda_k} (f(x_k) - f(x_{k+1})) \\
\Leftrightarrow 0 &\leq d^2(x_k, x_k) + \frac{2}{\lambda_k} (f(x_k) - f(x_{k+1})) \\
\Leftrightarrow \frac{2}{\lambda_k} (f(x_{k+1})) &\leq \frac{2}{\lambda_k} f(x_k) \\
\Leftrightarrow f(x_{k+1}) &\leq f(x_k)
\end{aligned}$$

Im nächsten Schritt dieses Beweises wollen wir nun zeigen, dass gilt:

$\lim_{k \rightarrow \infty} f(x_k) = f^* = \inf f(x)$. Diese Aussage zeigen wir über einen Widerspruchsbeweis, d.h. wir nehmen an dass $\lim_{k \rightarrow \infty} f(x_k) > f^*$. Für diese Fall existiere ein $x \in M$ und $\delta > 0$,

sodass $f(x) < f(x_k) \forall k$ gilt. Daraus folgt in Kombination mit der Ungleichung aus Lemma

4.5 folgendes:

$$d^2(x_{k+1}, x) \leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + \frac{2}{\lambda_k} (f(x_k) - f(x_{k+1}))$$

$$d^2(x_{k+1}, x) \leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + \frac{2}{\lambda_k} ((f(x_k) - \delta) - f(x_{k+1}))$$

$$d^2(x_{k+1}, x) \leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + \frac{2}{\lambda_k} f(x_k) - \frac{2\delta}{\lambda_k} - \frac{2}{\lambda_k} f(x_{k+1})$$

Da wir unter anderem bereits wissen, dass $f(x_{k+1}) \leq f(x_k)$ gilt nach weiteren Umformungen

weiter:

$$d^2(x_{k+1}, x) \leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + \frac{2\delta}{\lambda_k}$$

$$d^2(x_{k+1}, x) \leq d^2(x_k, x) + \frac{2\delta}{\lambda_k}$$

Nimmt man ausgehend von dieser Ungleichung weitere Abschätzungen vor sieht dies wie folgt aus:

$$d^2(x_{k+1}, x) \leq d^2(x_0, x) + \frac{2\delta}{\lambda_k}$$

$$2\delta \frac{1}{\lambda_k} \leq d^2(x_0, x) - d^2(x_{k+1}, x)$$

$$\frac{1}{\lambda_k} \leq \frac{1}{2\delta} (d^2(x_0, x) - d^2(x_{k+1}, x))$$

$$\Rightarrow \sum_{k=0}^j \frac{1}{\lambda_k} \leq \frac{1}{2\delta} (d^2(x_0, x) - d^2(x_{j+1}, x)) < \frac{1}{2\delta} d^2(x_0, x)$$

Da $\sum \frac{1}{\lambda_k} = +\infty$ sein soll folgt somit mit der letzten Zeile der Abschätzung der gewünschte

Widerspruch für den Fall, dass $\lim_{k \rightarrow \infty} f(x_k) > f^*$ und somit gilt

$$\lim_{k \rightarrow \infty} f(x_k) = f^* = \inf f(x).$$

Für den Beweis des zweiten Teils der Aussage sei U^* eine nichtleere Menge mit $\bar{x} \in U^*$

und es gilt $f(\bar{x}) \leq f(x_k) \forall k$. Betrachten wir nun erneut die Ungleichung aus Lemma 4.5 in

Kombination mit der soeben gegebenen Ungleichung, so erhalten wir: $d^2(x_{k+1}, \bar{x}) < d^2(x_k, \bar{x})$,

womit die Bedingungen für die Fejer-Konvergenz erfüllt sind und die Folge $\{x_k\}$ somit

Fejer-konvergent zu U^* ist. Im letzten Schritt dieses Beweises wollen wir zeigen, dass wenn ein Cluster-Punkt x^* der Folge $\{x_k\}$ in der nichtleeren Menge U^* enthalten ist, dass dann der

Limes der Folge genau dieser Cluster-Punkt ist. Hierzu verwenden wir die Aussage von

Lemma 4.4, aus welcher wir wissen, dass eine Fejer-konvergente Folge beschränkt ist.

Darüber hinaus wissen wir, dass M eine vollständige Mannigfaltigkeit ist und somit eine

konvergente Teilfolge zu $\{x_k\}$ existiert. Sei dann $\{x_{k_j}\}$ eine Teilfolge von $\{x_k\}$ für welche gilt: $\lim_{k \rightarrow \infty} f(x_{k_j}) = x^*$. Des Weiteren ist bekannt, dass die Funktion f stetig ist und das für diese, wie bereits weiter oben im Beweis gezeigt, gilt: $\lim_{k \rightarrow \infty} f(x_k) = f^*$. Hieraus folgt direkt, dass $x \in U^*$. Aufgrund dessen gehört der Cluster-Punkt x^* der Folge $\{x_k\}$ zur Menge U^* und laut Lemma 4.4 folgt direkt $\lim_{k \rightarrow \infty} x_k = x^*$. Damit haben wir alle Aussagen des Theorems bewiesen und die Konvergenz der Folge $\{x_k\}$, welche durch den Proximal-Punkt Algorithmus erzeugt wurde, gezeigt. □

Kapitel 5

Zusammenfassung

In dieser Bachelorarbeit haben wir einen Einblick in eine mögliche Optimierungsmethode auf Riemannschen Mannigfaltigkeiten gegeben.

Hierzu haben wir im ersten Kapitel zunächst einen Blick auf das Forschungsgebiet der Optimierungsmethoden geworfen, um uns eine Idee davon zu erschaffen, wofür die Optimierung in der Anwendung überhaupt benötigt wird. Dazu haben wir Beispiele genannt wie zum Beispiel Steuerungsprobleme in der Bildverarbeitung. Im nächsten Abschnitt haben wir dann kurz das Thema dieser Bachelorarbeit vorgestellt und kurz und knapp den Verlauf der Ausarbeitung beschrieben.

Im zweiten Kapitel stand die Darstellung des zugrunde liegenden Optimierungsproblems im Fokus. Dazu haben wir die Bedingungen genannt unter welchen die konvexe Funktion $f(x)$ minimiert werden soll. Dies geschah unter anderem unter Einbeziehung eines Startpunktes p_0 aus der Mannigfaltigkeit und einer Folge positiver Zahlen, welche mit $\{\lambda_k\}$ bezeichnet wurde. Daraufhin wurde die Iteration aufgeführt, durch welche den Proximal-Punkt Algorithmus die Folge $\{x_k\}$ generiert.

Im nächsten Kapitel stand die Darstellung der relevanten Aspekte aus den Bereichen der Riemannschen Geometrie sowie der konvexen Analysis im Vordergrund. Dieses Kapitel bildete die Grundlage für die weiteren Themen dieser Bachelorarbeit. Im Abschnitt, welcher sich mit den Grundlagen der Riemannschen Geometrie befasst, haben wir uns mit den Definitionen der Riemannschen Mannigfaltigkeit befasst sowie mit deren Beschaffenheiten. Hier haben wir uns unter anderem mit der Riemannschen Metrik, der Riemannschen Distanz und deren Vergleichbarkeit mit dem euklidischen Raum \mathfrak{R}^n befasst. Als wir näher auf die Riemannsche Metrik eingegangen sind, kamen Begriffe wie Längenfunktional, Winkel und Riemannsche Distanz mit ins Spiel. Mit Hilfe dieser Begrifflichkeiten kamen wir anschließend auf die Krümmung von Kurven, in wessen Zusammenhang der Begriff der Geodäte gefallen ist. Des Weiteren haben wir festgestellt, dass die Ableitung auf Mannigfaltigkeiten nicht so einfach zu bilden sind wie im euklidischen Raum, weshalb wir uns mit der Levi-Civita-Ableitung beschäftigt haben. Diese gleicht die Problematik der Vektoren in

unterschiedlichen Vektorräumen aus, sodass auch auf Mannigfaltigkeiten Ableitungen gebildet werden können. Ein weiterer Aspekt, mit dem wir uns in den Grundlagen befasst haben, ist die Vollständigkeit von riemannschen Mannigfaltigkeiten. Hier spielt unter anderem das Theorem von Hopf- Rinow eine ausschlaggebende Rolle. Ab diesem Zeitpunkt sind wir von der Vollständigkeit aller verwendeten Mannigfaltigkeiten ausgegangen. Um im weiteren Verlauf des Kapitels den Begriff der Hadamard-Mannigfaltigkeit definieren zu können, sind wir auf die Begrifflichkeiten des Krümmungstensors und der Schnittkrümmung eingegangen. Anhand dieser Definitionen konnten wir im Anschluss die Hadamard-Mannigfaltigkeit als vollständige, einfach zusammenhängende Mannigfaltigkeit mit nicht positiver Schnittkrümmung definieren. Im letzten Abschnitt der Einführung in die Riemannsche Geometrie haben wir noch den Begriff den geodätischen Dreiecks definiert, sowie ein relevantes Theorem eingeführt, welches den Zusammenhang zwischen den Ecken, Seiten und Winkeln des geodätischen Dreiecks hergestellt hat.

Im zweiten Teil des einführenden Kapitel haben wir uns mit der konvexen Analysis beschäftigt - genauer gesagt mit konvexen Mengen und konvexen Funktionen.

Hierzu haben wir zunächst sowohl den Begriff der konvexen Menge als auch den Begriff der konvexen Funktion definiert als auch bildlich dargestellt. Ferner haben wir zu beiden Themengebieten die wichtigsten Eigenschaften dargestellt, um uns einen Überblick über die wichtigsten Aspekte zu verschaffen. Im Abschnitt über konvexe Mengen haben wir uns weiter mit Aspekt befasst, dass es einen Punkt in unserer abgeschlossenen, konvexen Teilmenge C gibt, welcher den Abstand zu unserem $x' \in C$ minimiert. Um dies zu zeigen, haben wir einige Propositionen bewiesen. Abschließend haben wir in diesem Abschnitt noch einen Unterstützungsraum eingeführt. Dieser Unterraum wird als Support bezeichnet.

Im Abschnitt, in dem wir uns mit den konvexen Funktionen befasst haben, haben unter anderem die Begriffe Subgradient und Subdifferential eine Rolle gespielt.

Das darauf folgende Unterkapitel 3.3 hatte die Regularisierung zum Thema. Nachdem wir eine kurze Einführung zum allgemeinen Nutzen einer Regularisierung vorgenommen haben, sind wir direkt auf die Moreau-Yosida- Regularisierung eingegangen und haben diese definiert. Im Anschluss haben wir die Wohldefiniertheit dieser Regularisierung bewiesen. Hierfür mussten wir zuvor noch den Begriff der Koerzivität klären.

Das vierte Kapitel dieser Bachelorarbeit befasst sich mit dem Hauptthema dieser Ausarbeitung - dem Proximal-Punkt Algorithmus. Nachdem wir den Algorithmus vorgestellt haben, haben wir die Wohldefiniertheit bewiesen, welche eine direkte Folgerung aus der Wohldefiniertheit der Moreau-Yosida-Regularisierung darstellt.

Im abschließenden Unterkapitel haben wir die Konvergenz des Verfahrens gezeigt. Bei der Konvergenz des Verfahrens sprechen wir von der Fejer-Konvergenz, welche wir im ersten Schritt definiert haben. Die beiden darauf folgenden Lemma, welche wir bewiesen haben, vervollständigten den Beweis der Konvergenz der durch den Proximal-Punkt Algorithmus erzeugten Folge.

Abschließend lässt sich sagen, dass der Proximal-Punkt Algorithmus ein wohldefiniertes und konvergentes Verfahren ist, um Optimierungsprobleme der dargestellten Art zu lösen.

Bild- und Literaturverzeichnis

1. O.P. Ferreira & P.R. Oliveira - Artikel: "*Proximal Point Algorithm on Riemannian Manifolds*" 20.08.2008
2. Dirk Ferus - Skript "Riemannsche Geometrie" SoSe 2004 TU Berlin-
<https://page.math.tu-berlin.de/~ferus/DG/RiemGeom.pdf>
3. Patrick Krause - *Einführung in die Optimierungsanwendung* -Skript zum Seminar "Optimierung und Variationsmethoden in der Bildverarbeitung und den Naturwissenschaften" WS 2017/2018 WWU Münster
4. Robert E. Mahony & Rodolphe Sepulchre - "*Optimization Algorithm on Matrix Manifolds*" 18.07.2007
5. Jan-Hendrik Treude - Skript zur Vorlesung "Riemannsche Geometrie" aus dem Jahr 2019 an der Uni Konstanz
http://www.math.uni-konstanz.de/~treude/Dateien/RG_Skript.pdf
6. Franz Schuster – Skript zur Vorlesung "Harmonische Analysis und Geometrie" aus dem Wintersemester 2019 <https://dmg.tuwien.ac.at/schuster/SkriptHarmAnaGeo.pdf>
7. Kapitel 3 aus dem Vorlesungsskript der Wias Berlin - "Konvexität"
https://www.wias-berlin.de/people/john/LEHRE/OPTI/SS07/opti_nlin03.pdf
8. Jürgen Tietze - "Einführung in die angewandte Wirtschaftsmathematik" - Anhang A
<https://link.springer.com/content/pdf/bbm:978-3-663-01968-8/1.pdf>
9. Wilking - Skript zur Vorlesung Differentialgeometrie - Universität Münster
<https://ivv5hpp.uni-muenster.de/u/wilking/Diffgeo1/diffgeo1.pdf>
10. T. Sakai - 1996-Riemannian Geometry, translation of mathematical monographs
11. Abbildungen:
https://favpng.com/png_view/mathematics-epigraph-convex-function-convex-set-mathematics-png/EBGTJn88
12. Abbildung: Segmentierung in der medizinischen Bildverarbeitung-
<https://docplayer.org/69847326-Segmentierung-in-der-medizinischen-bildverarbeitung.html>

Eigenständigkeitserklärung

Hiermit versichere ich, dass die vorliegende Arbeit über "Proximal-Punkt Algorithmen auf Riemannschen Mannigfaltigkeiten" selbstständig von mir und ohne fremde Hilfe verfasst worden ist, dass keine anderen Quellen und Hilfsmittel als die angegebenen benutzt worden sind und dass die Stellen der Arbeit, die anderen Werken – auch elektronischen Medien – dem Wortlaut oder Sinn nach entnommen wurden, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht worden sind. Mir ist bekannt, dass es sich bei einem Plagiat um eine Täuschung handelt, die gemäß der Prüfungsordnung sanktioniert werden kann.

Ich erkläre mich mit einem Abgleich der Arbeit mit anderen Texten zwecks Auffindung von Übereinstimmungen sowie mit einer zu diesem Zweck vorzunehmenden Speicherung der Arbeit in einer Datenbank einverstanden.

Ich versichere, dass ich die vorliegende Arbeit oder Teile daraus nicht anderweitig als Prüfungsarbeit eingereicht habe.

Münster, 29.09.2022