

Kapitel 1

Einführung

Seit langem beschäftigt sich die Mathematik mit dem Auffinden extremaler Werte von Funktionen und der Lage der entsprechenden Extremalstellen. Neben dem rein theoretischen Aspekt war diese Entwicklung immer auch getrieben von konkreten Anforderungen, etwa aus der Physik oder der Ökonomie, bei der nach Parametern gesucht wurde, für die der Wert einer konstruierten Größe einen maximalen oder minimalen Betrag erreicht. Die mathematische Disziplin, die sich mit derartigen Aufgaben beschäftigt, bezeichnen wir als Optimierung.

Häufig verlangt die Situation vom Anwender aber auch, daß nicht nur eine solche Forderung, sondern mehrere Bewertungskriterien bestens erfüllt werden sollen. Hierbei spricht man von Problemen der multikriteriellen oder Mehrziel- oder auch Vektoroptimierung.

Das Wesen solcher Probleme besteht darin, daß es für gewöhnlich keine Lösungen gibt, die alle Ziele gleichzeitig zum Optimum führen. Vielmehr ist es so, daß sich die Zielkriterien stark widersprechen, so daß der Optimalpunkt der einen Zielfunktion sehr schlecht in Bezug auf ein anderes Kriterium ist. (Ein Problem, bei dem die Ziele simultan erreicht werden können, ist auch ohne die Methoden der Mehrzieloptimierung zu lösen. Solche Aufgaben seien daher für unsere Zwecke nicht „wesentlich“.)

Den Ausweg aus dieser Lage liefert eine Vorgehensweise, die auch außerhalb der Wissenschaften zur Schlichtung in Streitfällen gern strapaziert wird: Wenn es schon keine Möglichkeit gibt, allen Anforderungen gerecht zu werden, dann soll wenigstens eine Lösung gefunden werden, die einen Kompromiß zwischen den sich widersprechenden Zielen liefert. Eine solche Lösung soll so beschaffen sein, daß in einem noch festzulegenden Sinne alle Ziele gleichzeitig möglichst gut erfüllt werden. In der Vektoroptimierung bezeichnet man solche Kompromißlösungen auch als *effiziente* Punkte.

Eine sehr populäre, weil auch naheliegende, Definition von Kompromißlösungen stammt von dem italienischen Volkswirtschaftler und Soziologen V. Pareto (1848–1923). Demnach wird eine Alternative von einer anderen *dominiert*, wenn die letztere bezüglich aller Ziele nicht schlechter und hinsichtlich mindestens eines Zieles sogar besser als die erste ist. Eine Alternative, die von keiner anderen erlaubten Möglichkeit dominiert wird, heißt dann effizient.

Diese Begriffe sollen am folgenden Beispiel verdeutlicht werden. Weiterhin zeigt dieses Beispiel, daß auch im täglichen Leben ständig multikriterielle Entscheidungsfindungen vorgenommen werden müssen.

Beispiel 1.1 *Ein Reisender sucht eine optimale Zugverbindung für die Fahrt von Halle (S) Hbf. nach Passau Hbf. In Betracht kommen alle Züge, die Halle zwischen 7.00 Uhr und 10.00 Uhr verlassen. Er stellt an die „Güte“ der Verbindung die folgenden Anforderungen.*

1. Die Abfahrtszeit AB liegt möglichst spät.
2. Die Ankunftszeit AN liegt möglichst früh.
3. Die Zahl der Umstiege UM sei möglichst gering.
4. Der Fahrpreis PR sei möglichst gering.

Ein Blick in den Fahrplan liefert ihm diese Alternativen:

Nr.	Abfahrt (AB)	Ankunft (AN)	Fahrzeit (FZ)	umsteigen (UM)	benutzte Züge	Strecke (ST)	Preis (PR)
1	7.32	14.34	7:02 h	1	IR, ICE	706 km	209,00 DM
2	7.32	14.34	7:02 h	2	IR, EC, ICE	507 km	155,00 DM
3	7.58	14.34	6:36 h	2	SE, IR, ICE	518 km	146,00 DM
4	8.31	20.39	12:08 h	0	IC	1230 km	342,00 DM
5	9.32	15.50	6:18 h	2	IR, IC, IC	507 km	144,00 DM

Quelle: Deutsche Bahn AG, Winterfahrplan 98/99.

Die Verbindung 2 wird als von der Alternative 3 dominiert. Letztere hat eine spätere Abfahrtszeit, erfordert die selbe Zahl an Umstiegen und ist darüberhinaus nicht teurer als erstere. Die Verbindungen 1,3,4 und 5 hingegen sind effizient. Selbst die offenbar absurde Reisemöglichkeit 4 (über Köln) ist in der Hinsicht eine effiziente Alternative, als daß es keine schnellere oder billigere Verbindung gibt, die ebenfalls ohne Umsteigen auskommt.

Beispiel 1.2 Im vorangegangenen Beispiel werden die Kriterien 1 und 2 durch die folgende Forderung ersetzt.

Die Fahrzeit $FZ = AN - AB$ sei möglichst kurz.

Wie im Beispiel 1.1 wird 2 von 3 dominiert. Desweiteren dominiert die Alternative 5 die Verbindung 3 und mithin auch 2. Effiziente Alternativen sind 1,4 und 5.

Beispiel 1.3 Um sich die Entscheidung zu erleichtern, entschließt sich der Reisende aus den beiden vorangegangenen Beispiele, die von ihm aufgestellten Anforderungen zu bewerten. Unterstellen wir ihm eine materialistische Denkweise und sagen, er ordne jedem dieser Kriterien einen Geldwert zu. Eine im Zug verbrachte Stunde ist ihm 100,00 DM wert. Die Kosten für einen Umstieg beziffert er mit 20,00 DM. Das einzige Kriterium für die Qualität einer Verbindung lautet dann:

Der Gesamtpreis $GP = 100FZ + 20UM + PR$ sei möglichst gering.

Aus diesen Vorgaben errechnet man:

Nr. Nr.	Kosten der Fahrzeit ($100FZ$)	Kosten des Umsteigens ($20UM$)	Kosten des Fahrscheins ($1PR$)	Gesamtpreis (GP)
1	703,33 DM	20,00 DM	209,00 DM	932,33 DM
2	703,33 DM	40,00 DM	155,00 DM	898,33 DM
3	660,00 DM	40,00 DM	146,00 DM	846,00 DM
4	1213,33 DM	0,00 DM	349,00 DM	1562,33 DM
5	630,00 DM	40,00 DM	144,00 DM	814,00 DM

Offensichtlich ist unter dieser Bewertung die Verbindung 5 die günstigste. Die auf diese Weise ermittelte Lösung ist im übrigen eine der effizienten Alternativen von Beispiel 1.2.

Die im Beispiel 1.3 vorgenommene „Bewertung“ ist unabhängig von dem hier betrachteten Problem eine oft verwendete Methode zur Lösung multikriterieller Aufgaben. Aus mathematischer Sicht wurde dabei aus dem Vektor, bestehend aus den drei Zielen (FZ, UM, PR) und dem Vektor der Bewertungen ($100, 20, 1$) das Skalarprodukt

$$\left\langle \begin{pmatrix} 100 \\ 20 \\ 1 \end{pmatrix}, \begin{pmatrix} FZ \\ UM \\ PR \end{pmatrix} \right\rangle = 100FZ + 20UM + PR$$

gebildet und dieses minimiert. Eine solche Vorgehensweise, bei der anstelle eines Vektors von Zielen eine daraus gebildete skalare Größe optimiert wird, heißt *Skalarisierung*.

Wie wir später zeigen werden, ist für bestimmte Klassen von Skalarisierungen die Lösung der skalarisierten Aufgabe stets effizient für das vektorielle Optimierungsproblem. Im vorliegenden Fall hätte die Bewertung von FZ , UM und PR mit einem beliebigen Preis größer als Null immer eine der in Beispiel 1.2 als effizient ermittelte Alternativen ergeben.

Wir wollen nun die bis hierher intuitiv eingeführten Begriffe in eine exakte mathematische Sprechweise übertragen. Dazu bedarf es zunächst einiger grundlegender Festlegungen.

Wir bezeichnen eine Teilmenge M eines linearen Raumes \mathcal{X} als *konvex*, wenn für je zwei Punkte $x, y \in M$ gilt

$$\lambda x + (1 - \lambda)y \in M \quad \forall \lambda \in [0, 1]. \quad (1.1)$$

Mit $\text{conv}(M)$ beschreiben wir die konvexe Hülle der Menge M , also

$$\text{conv}(M) := \{z \in \mathcal{X} \mid \exists x, y \in M; \exists \lambda \in [0, 1] : z = \lambda x + (1 - \lambda)y\}. \quad (1.2)$$

Für Teilmengen M eines topologischen Raumes verstehen wir unter $\text{int } M$ das Innere, unter $\text{cl } M$ den Abschluß sowie unter $\text{bd } M$ den Rand der Menge.

Es sei R^p der p -dimensionale euklidische Raum aller p -Tupel $x = (x_1, \dots, x_p)$. Zur Verdeutlichung bezeichnen wir das Nullelement des R^p in allen Zweifelsfällen mit 0^p .

Der Raum R^p ist mit der Norm

$$\|x\| := \sqrt{\sum_{i=1}^p (x_i)^2} \quad \forall x \in R^p \quad (1.3)$$

und dem Skalarprodukt

$$\langle x, y \rangle := \sqrt{\sum_{i=1}^p x_i y_i} \quad \forall x, y \in R^p \quad (1.4)$$

versehen.

Eine Teilmenge $K \subseteq R^p$ heißt *Kegel*, wenn für jeden Punkt $x \in K$ gilt

$$\alpha x \in K \quad \forall \alpha \geq 0. \quad (1.5)$$

Mit $\text{cone}(M)$ bezeichnen wir den Kegel, der durch die Menge M aufgespannt wird, das heißt

$$\text{cone}(M) := \{y \in R^p \mid \exists \alpha \geq 0, \exists x \in M : y = \alpha x\}. \quad (1.6)$$

Mit $\text{convcone}(M)$ bezeichnen wir den durch M aufgespannten konvexen Kegel

$$\text{convcone}(M) := \left\{ y \in R^p \mid \exists k, \exists \alpha_1, \dots, \alpha_k \geq 0, \exists x^1, \dots, x^k \in M : y = \sum_{i=1}^k \alpha_i x^i \right\}. \quad (1.7)$$

Zu einem Kegel $K \subseteq R^p$ wird mit

$$K^* := \{x \in R^p \mid \langle x, y \rangle \geq 0 \forall y \in K\} \quad (1.8)$$

der zugehörige *duale Kegel* definiert.

Ein häufig betrachteter Kegel im Raum R^p ist der sogenannte positive Orthant

$$R_+^p := \{y \in R^p \mid y_i \geq 0 \ (i = 1, \dots, p)\}.$$

Eine reellwertige Funktion ϕ , definiert auf einer konvexen Menge M , bezeichnen wir als konvex, wenn für alle $x, y \in M$ gilt

$$\phi(\lambda x + (1 - \lambda)y) \leq \lambda\phi(x) + (1 - \lambda)\phi(y) \quad \forall \lambda \in [0, 1]. \quad (1.9)$$

Vektorwertige Funktionen $\Phi : M \rightarrow R^p$ werden konvex genannt, wenn für $x, y \in M$ gilt

$$\Phi(\lambda x + (1 - \lambda)y) \in \lambda\Phi(x) + (1 - \lambda)\Phi(y) - R_+^p \quad \forall \lambda \in [0, 1]. \quad (1.10)$$

Dies bedeutet lediglich, daß $\Phi = (\Phi_1, \dots, \Phi_p)$ genau dann konvex ist, wenn jede Komponente Φ_i ($i = 1, \dots, p$) von Φ eine reellwertige konvexe Funktion ist.

Eine reellwertige Funktion ϕ , definiert auf einem linearen normierten Raum \mathcal{X} , heißt Fréchet-differenzierbar im Punkt $x_0 \in \mathcal{X}$, wenn es auf \mathcal{X} ein lineares, stetiges Funktional $\phi'(x_0)$ gibt mit

$$\phi(x_0 + x) = \phi(x_0) + \phi'(x_0)(x) + \|x\|s(x) \quad \forall x \in \mathcal{X}, \quad (1.11)$$

wobei $s(x)$ der Bedingung $\lim_{x \rightarrow 0} s(x) = 0$ genügt. Das Funktional $\phi'(x_0)$ heißt Fréchet-Ableitung von ϕ im Punkt x_0 . Analog definieren wir die Fréchet-Ableitung vektorwertiger Funktionen $\Phi : \mathcal{X} \rightarrow R^p$ durch

$$\Phi(x_0 + x) = \Phi(x_0) + \Phi'(x_0)(x) + \|x\|S(x) \quad \forall x \in \mathcal{X}, \quad (1.12)$$

Hierbei verwenden wir als Ableitung eine lineare, stetige Funktion $\Phi'(x_0) : \mathcal{X} \rightarrow R^p$. Für das Restglied $S(x) \in R^p$ gelte entsprechend $\lim_{x \rightarrow 0} S(x) = 0^p$.

Für lineare, stetige Funktionale ϕ auf einem linearen Raum mit der Norm $\|\cdot\|$ definieren wir die *duale Norm* durch

$$\|\phi\|_* := \sup_{\|x\|=1} |\phi(x)|. \quad (1.13)$$

Mit den so eingeführten Bezeichnungen wollen den Effizienzbegriff von Pareto konkretisieren: Ein Punkt $y \in R^p$ wird von $x \in R^p$ *dominiert*, wenn

$$x \in y - R_+^p \setminus \{0^p\}. \quad (1.14)$$

Ein Punkt $y \in Y$ einer Teilmenge $Y \subseteq R^p$ heißt *Pareto-effizient für die Menge Y*, wenn er von keinem anderen Punkt aus Y dominiert wird, das heißt

$$(y - R_+^p \setminus \{0^p\}) \cap Y = \emptyset. \quad (1.15)$$

Durch den Kegel R_+^p wird im Raum R^p eine Vergleichsrelation eingeführt. Diese wird auch als die „natürliche“ Vergleichsrelation bezeichnet, weil hier die Vektoren in natürlicher Weise, das heißt komponentenweise verglichen werden. Eine allgemeinere Formulierung der Effizienz wird erreicht, indem wir eine solche Relation durch einen beliebigen konvexen Kegel $K \subseteq R^p$ festlegen.

Ein Punkt $y \in R^p$ wird von $x \in R^p$ *bezüglich K dominiert*, wenn

$$x \in y - K \setminus \{0^p\}. \quad (1.16)$$

In dieser Sprechweise heißt ein Punkt $y \in Y$ *effizient für die Menge Y bezüglich K*, wenn

$$(y - K \setminus \{0^p\}) \cap Y = \emptyset. \quad (1.17)$$

Offenbar ist die Pareto-Effizienz ein Spezialfall der so eingeführten Effizienz bezüglich eines Kegels, wenn man sich davon überzeugt, daß R_+^p ein konvexer Kegel ist.

Bemerkung 1.1 *Ist der konvexe Kegel K darüberhinaus spitz, das heißt*

$$K \cap (-K) = \{0^p\},$$

so ist die durch K festgelegte Vergleichsrelation \leq_K , definiert als

$$x \leq_K y \iff y - x \in K,$$

reflexiv, transitiv und antisymmetrisch, mithin eine Halbordnung. Wir können in dieser Arbeit auf die Spitzheit der Ordnungskegel verzichten. Man beachte, daß dadurch die Antisymmetrie der Vergleichsrelation im allgemeinen nicht gegeben ist.

Die „natürliche“ Vergleichsrelation ist eine Halbordnung, da R_+^p ein spitzer konvexer Kegel ist.

Eine Abschwächung des letzteren Begriffs stellt die schwache Effizienz dar: Es sei $K \subseteq R^p$ ein konvexer Kegel mit nichtleerem Inneren. Ein Punkt $y \in Y \subseteq R^p$ heißt *schwach effizient für die Menge Y bezüglich K* , wenn

$$(y - \text{int } K) \cap Y = \emptyset. \quad (1.18)$$

Tatsächlich ist jeder effiziente Punkt einer Menge auch schwach effizient für diese.

Mit dem anfangs angeführten Beispiel 1.3 haben wir die Methode der Skalarisierung multikriterieller Optimierungsaufgaben vorgestellt. Hierbei wird anstelle eines Problems

Bestimme alle Punkte $y \in Y$, die effizient für die Menge Y bezüglich K sind!

das Problem

Bestimme alle Punkte $y \in Y$, die den Ausdruck $\langle \lambda, y \rangle$ über Y minimieren!

mit einem geeignet gewählten Skalarisierungsvektor $\lambda \in K^*$ betrachtet.

Am Beispiel zeigte sich, daß mit der dort gewählten Skalarisierung der drei Kriterien eine Lösung gefunden wurde, die auch ein effizienter Punkt der entsprechenden nichtskalarisierten Aufgabe war. Dies ist in der Tat ein allgemeines Phänomen.

Satz 1.1 *Sei Y eine nichtleere Teilmenge von R^p und K ein konvexer Kegel in R^p . Gegeben sei ein Funktional $\lambda \in K^*$.*

Dann ist jeder Punkt $y_0 \in Y$ mit

$$\langle \lambda, y_0 \rangle \leq \langle \lambda, y \rangle \quad \forall y \in Y \quad (1.19)$$

a) *schwach effizient für die Menge Y bezüglich K , falls $\lambda \neq 0^p$;*

b) *effizient für die Menge Y bezüglich K , falls λ im Quasi-Inneren von K^* liegt, das heißt*

$$\langle \lambda, y \rangle > 0 \quad \forall y \in K \setminus \{0^p\}.$$

Beweis. a) Angenommen, y_0 wäre nicht schwach effizient für die Menge Y bezüglich K . Dann gibt es einen Punkt $y_1 \in Y$ mit

$$y_1 \in y_0 - \text{int } K.$$

Weil für $\lambda \neq 0^p$ gilt

$$\langle \lambda, y \rangle > 0 \quad \forall y \in \text{int } K,$$

folgt hieraus

$$\langle \lambda, y_1 \rangle < \langle \lambda, y_0 \rangle,$$

was der Voraussetzung (1.19) widerspricht.

b) Angenommen, y_0 wäre nicht effizient für die Menge Y bezüglich K . Dann gibt es einen Punkt $y_1 \in Y$ mit

$$y_1 \in y_0 - K \setminus \{0^p\}.$$

Weil λ im Quasi-Inneren von K^* liegt, folgt hieraus

$$\langle \lambda, y_1 \rangle < \langle \lambda, y_0 \rangle,$$

was der Voraussetzung (1.19) widerspricht. ■

Von besonderem Interesse ist allerdings die Umkehrung dieses Satzes. Es stellt sich die Frage, ob alle effizienten Punkte eines multikriteriellen Optimierungsproblems auch Lösung eines entsprechenden skalarisierten Problems sind, wenn man den Multiplikator λ nur hinreichend oft variiert. Gelingt es, dieses zu zeigen, so können wir die Gesamtheit aller effizienten Punkte durch Lösung einer Schar von skalaren Aufgaben finden, wofür uns die Methoden der gewöhnlichen Optimierung zur Verfügung stehen.

Das unterstreicht die Bedeutung von Aussagen des folgenden Typs:

Sei $y_0 \in Y$ ein effizienter Punkt für die Menge Y bezüglich K . Dann gibt es (unter gewissen weiteren Voraussetzungen) ein $\lambda \in K^$, $\lambda \neq 0^p$, so daß y_0 den Ausdruck $\langle \lambda, y \rangle$ über Y minimiert.*

Im Verlauf dieser Arbeit werden desöfteren Resultate vorgestellt, die vom Prinzip her solchen „Skalarisierungsregeln“ gleichkommen.

Eine andere Vorgehensweise, auf die wir zurückgreifen werden, ist die unter dem Namen von J. de Lagrange (1736–1813) bekannt gewordenen Methode der Multiplikatorenregeln. Diese dienen zur Behandlung von Optimierungsproblemen, wenn die Bedingungen für die Zulässigkeit einer Lösung von spezieller funktionaler Gestalt sind. In ihrer ursprünglichen Form galten sie für die Optimierung einer reellen Funktion ϕ über einer Menge der Form

$$S = \{x \in X \mid \psi_1(x) = 0, \dots, \psi_k(x) = 0\}. \quad (1.20)$$

Unter gewissen Voraussetzungen wurde nachgewiesen, daß x_0 genau dann eine Extremalstelle von ϕ

über S ist, wenn auch die „Lagrange-Funktion“

$$L(\lambda, x) = \phi(x) + \sum_{i=1}^k \lambda_i \psi_i(x) \quad (1.21)$$

in x_0 ein Extremum über $R^k \times X$ annimmt. Dadurch war es möglich, anstelle eines restringierten ein freies Optimierungsproblem zu bearbeiten. Gelöst werden konnte dieses Problem mit Hilfe der Differentialrechnung, wenn man die Differenzierbarkeit aller auftretenden Funktionen unterstellte.

Im Verlauf weiterer Arbeiten wurde die Idee, die Zielfunktion auf die oben beschriebene Weise zu erweitern, in vielfältigen Formen angewandt. Zusätzlich zu den Gleichungsrestriktionen konnten bald auch Nebenbedingungen in Gestalt von Ungleichungen $\psi_i(x) \leq 0$ ($i = k + 1, \dots, m$) betrachtet werden, wenn man die entsprechenden Multiplikatoren $\lambda_{k+1}, \dots, \lambda_m$ als nichtnegativ voraussetzte. Es war desweiteren möglich, die in den Ungleichungen auftretende Vergleichsrelation mit Hilfe eines allgemeinen Ordnungskegels auszudrücken, so wie wir es bei der Verallgemeinerung des Effizienzbegriffes von Pareto vorgeführt hatten. Andererseits erlaubten es die Einwicklungen in der Analysis, die Optimalität einer Lösung auf andere Weise als durch die klassische Ableitung zu charakterisieren, wodurch die Voraussetzung der Differenzierbarkeit der Funktionen abgeschwächt werden konnte. Mit der Vielfalt der in der modernen Optimierungstheorie betrachteten Modelle und den unterschiedlichsten Charakterisierungsmöglichkeiten von Extremalität entstand eine solche Fülle von „Lagrangeschen Multiplikatorenregeln“, daß selbst ein Überblick über all diese einen sehr großen Aufwand erfordern würde.

Die in dieser Arbeit vorgestellten Optimalitätsbedingungen werden eine Kombination aus Skalarisierungs- und Multiplikatorenregeln, wie sie oben vorgestellt wurden, sein. Auch hierbei sollen auftretende Nebenbedingungen in funktionaler Form nach Art und Weise der Lagrangeschen Multiplikatorenregeln behandelt werden, indem wir anstelle der Zielfunktion eine die Nebenbedingungen umfassende Lagrange-Funktion betrachten.

Die Charakterisierung der Optimalität einer Lösung geschieht mit Hilfe sogenannter derivierter Mengen. Diese stellen einen recht allgemeinen Ableitungsbegriff dar. Im Kapitel 2 werden wir die entsprechenden Definitionen einführen. Ferner werde ich einen neuen Beweis für die grundlegende Multiplikatorenregel über derivierte Mengen in der multikriteriellen Optimierung vorstellen.

Im Kapitel 3 soll darauf eingegangen werden, daß derivierte Mengen oft in Form konvexer Kegel auftreten. Es liegt daher nahe, gewisse Beziehungen zu anderen Kegeln, die ebenfalls Eigenschaften einer Ableitung aufweisen, zu untersuchen.

Kapitel 4 ist der Betrachtung von Multiplikatorenregeln für Näherungslösungen von Optimierungsaufgaben gewidmet. Ich werde unter Verwendung der Variationsprinzips von Ekeland eine Multiplikatorenregel über derivierte Mengen für Näherungslösungen herleiten. Dieses allgemeine Ergebnis wird in Spezialfällen diskutiert.

Abschließend werden wir uns in Kapitel 5 mit Regularitätsbedingungen beschäftigen. Diese sichern, daß der Multiplikator, der in der Lagrange-Funktion die Zielfunktion bewertet, nicht Null wird. In der Literatur fand sich keine solche Aussage für die im Kapitel 2 betrachteten Multiplikatorenregeln. Ich habe solche Bedingungen hergeleitet. Bemerkenswert an diesen ist, daß die Regularität einer Lösung hierbei auch unter Einbeziehung der Zielfunktion festgestellt wird, während sich Regularitätsbedingungen für gewöhnlich ausschließlich auf Eigenschaften der zulässigen Menge stützen. Zwischen diesen und den von mir gefundenen Resultaten sollen Vergleiche gezogen werden.