Vector Optimization

Johannes Jahn

Vector Optimization

Theory, Applications, and Extensions

Second Edition

Springer

Prof. Dr. Johannes Jahn
Universität Erlangen-Nürnberg
Department Mathematik
Martensstraße 3
91058 Erlangen
Germany
jahn@am.uni-erlangen.de

ISBN 978-3-642-17004-1 e-ISBN 978-3-642-17005-8
DOI 10.1007/978-3-642-17005-8
Springer Heidelberg Dordrecht London New York

Cover design: WMXDesign GmbH

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To Claudia and Martin

Preface

In vector optimization one investigates optimal elements such as minimal, strongly minimal, properly minimal or weakly minimal elements of a nonempty subset of a partially ordered linear space. The problem of determining at least one of these optimal elements, if they exist at all, is also called a vector optimization problem. Problems of this type can be found not only in mathematics but also in engineering and economics. Vector optimization problems arise, for example, in functional analysis (the Hahn-Banach theorem, the Bishop-Phelps lemma, Ekeland's variational principle), multiobjective programming, multi-criteria decision making, statistics (Bayes solutions, theory of tests, minimal covariance matrices), approximation theory (location theory, simultaneous approximation, solution of boundary value problems) and cooperative game theory (cooperative n player differential games and, as a special case, optimal control problems). In the last two decades vector optimization has been extended to problems with set-valued maps. This new field of research, called set optimization, seems to have important applications to variational inequalities and optimization problems with multivalued data.

The roots of vector optimization go back to F.Y. Edgeworth (1881) and V. Pareto (1906) who have already given the definition of the standard optimality concept in multiobjective optimization. But in mathematics this branch of optimization has started with the legendary paper of H.W. Kuhn and A.W. Tucker (1951). Since about the end of the 1960's research is intensively made in vector optimization.

It is the aim of this book to present various basic and important results of vector optimization in a general mathematical setting and to demonstrate its usefulness in mathematics and engineering. An extension to set optimization is also given. The first three parts are a revised edition of the former book [160] of the author. The forth part on engineering applications and the fifth part entitled extensions to set optimization have been added.

The theoretical vector optimization results are contained in the second part of this book. For a better understanding of the proofs several theorems of convex analysis are recalled in the first part. This part concisely summarizes the necessary background material and may be viewed as an appendix.

The main part of this book begins on page 102 with a discussion of several optimality notions together with some simple relations. Necessary and sufficient conditions for optimal elements are obtained by scalarization, i.e. the original vector optimization problem is replaced by an optimization problem with a real-valued objective map. The scalarizing functionals being used are certain linear functionals and norms. Existence theorems for optimal elements are proved using Zorn's lemma and the scalarization theory. For vector optimization problems with inequality and equality constraints a generalized Lagrange multiplier rule is given. Moreover, a duality theory is developed for convex maps. These results are also specialized to abstract linear optimization problems. The third part of this book is devoted to the application of the preceding general theory. For vector approximation problems the connections to simultaneous approximation problems are shown and a generalized Kolmogorov condition is formulated. Furthermore, nonlinear and linear Chebyshev problems are considered in detail. The last section is entitled cooperative n player differential games. These include optimal control problems. For these games a maximum principle is proved.

In the part on engineering applications the developed theoretical results are applied to multiobjective optimization problems arising in engineering. After a presentation of the theoretical basics of multiobjective optimization numerical methods are discussed. Some of these methods are applied to concrete nonlinear multiobjective optimization problems from electrical engineering, computer science, chemical engineering and medical engineering. The last part extends the second part of this book to set optimization. After an introduction to this field of research including basic concepts the notion of the contingent epiderivative is discussed in detail. Subdifferentials are the topic together with a comprehensive chapter on optimality conditions in set optimization.

This book should be readable for students in mathematics whose background includes a basic knowledge in optimization and linear functional analysis. Mathematically oriented engineers may be interested in the forth part on engineering applications.

The bibliography contains only a selection of references. A reader who is interested in the first papers of vector optimization is requested to consult the extensive older bibliographies of Achilles-Elster-Nehse [1], Nehse [258] and Stadler [312].

This second edition is a revised version containing two new sections, additional remarks on the contribution of Edgeworth and Pareto and an updated bibliography.

I am very grateful to Professors W. Krabs, R.H. Martin and B. Brosowski for their support and valuable suggestions. I also thank Dr. D. Diehl, Dr. G. Eichfelder and Dr. E. Schneider for useful comments. Moreover, I am indebted to A. Garhammer, S. Gmeiner, Dr. J. Klose, Dr. A. Merkel, Dr. B. Pfeiffer and H. Winkler for their assistance.

Erlangen, September 2010 Johannes Jahn

Contents

Part I

Convex Analysis

Convex analysis turns out to be a powerful tool for the investigation of vector optimization problems in a partially ordered linear space for two main reasons. A partial ordering in a real linear space can be characterized by a convex cone and, therefore, theorems concerning convex cones are very useful. Furthermore, separation theorems are especially helpful for the development of a Lagrangian theory. In this first part which consists of three chapters we present all these results on convex analysis which are necessary for the following theory on vector optimization. The most important theorems are separation theorems, a James theorem and a Krein-Rutman theorem.

Chapter 1

Linear Spaces

Although several results of the theory described in the second part of this book are also valid in a rather abstract setting we restrict our attention to real linear spaces. For convenience, we summarize in this chapter the well-known definitions of linear spaces and convex sets as well as the definition of (locally convex) topological linear spaces and we consider a partial ordering in such a linear setting. Finally, we investigate some special partially ordered linear spaces and list various known properties.

1.1 Linear Spaces and Convex Sets

We recall the definition of a real linear space and present some other notations.

Definition 1.1. Let X be a given set. Assume that an addition on X, i.e. a map from $X \times X$ to X, and a scalar multiplication on X, i.e. a map from $\mathbb{R} \times X$ to X, is defined. The set X is called a *real linear space*, if the following axioms are satisfied (for arbitrary $x, y, z \in X$ and $\lambda, \mu \in \mathbb{R}$):

(a) $(x + y) + z = x + (y + z)$,

(b) $x + y = y + x$,

(c) there is an element $0_X \in X$ with $x + 0_X = x$ for all $x \in X$,

J. Jahn, *Vector Optimization: Theory, Applications, and Extensions*,
DOI 10.1007/978-3-642-17005-8_1, © Springer-Verlag Berlin Heidelberg 2011

(d) for every $x \in X$ there is a $y \in X$ with $x + y = 0_X$,

(e) $\lambda(x + y) = \lambda x + \lambda y$,

(f) $(\lambda + \mu)x = \lambda x + \mu x$,

(g) $\lambda(\mu x) = (\lambda\mu)x$,

(h) $1x = x$.

The element 0_X given under (c) is called the *zero element* of X.

Definition 1.2. Let S and T be nonempty subsets of a real linear space X. Then we define the *algebraic sum* of S and T as

$$S + T := \{x + y \mid x \in S \text{ and } y \in T\}$$

and the *algebraic difference* of S and T as

$$S - T := \{x - y \mid x \in S \text{ and } y \in T\}.$$

For an arbitrary $\lambda \in \mathbb{R}$ the notation λS will be used as

$$\lambda S := \{\lambda x \mid x \in S\}.$$

It is important to note that the set equation $S + S = 2S$ does not hold in general for a nonempty subset S of a real linear space.

Definition 1.3. Let X be a real linear space. The set X' is defined to be the set of all linear maps from X to $\mathbb{R}$. If we define for all $\varphi, \psi \in X'$ and all $\lambda \in \mathbb{R}$

$$(\varphi + \psi)(x) = \varphi(x) + \psi(x) \quad \text{for all } x \in X$$

and

$$(\lambda\varphi)(x) = \lambda\,\varphi(x) \quad \text{for all } x \in X,$$

then X' is a real linear space itself and it is called the *algebraic dual space* of X. The algebraic dual space of X' is denoted by X'' and it is called the *second algebraic dual space* of X.

The most important class of subsets in a real linear space are convex sets.

Definition 1.4. Let S be a subset of a real linear space X.

(a) Let some $\bar{x} \in S$ be given. The set S is called *starshaped* at $\bar{x}$, if for every $x \in S$

$$\lambda x + (1 - \lambda)\bar{x} \in S \quad \text{for all } \lambda \in [0, 1]$$

(see Fig. 1.1).

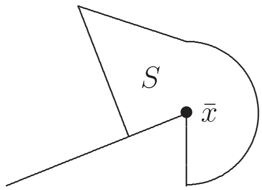

Figure 1.1: A set S being starshaped at $\bar{x}$.

(b) The set S is called *convex*, if for every $x, y \in S$

$$\lambda x + (1 - \lambda)y \in S \quad \text{for all } \lambda \in [0, 1]$$

(see Fig. 1.2 and 1.3).

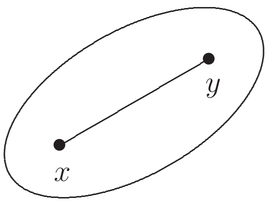

Figure 1.2: Convex set.

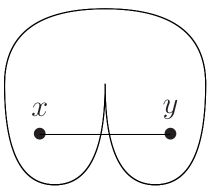

Figure 1.3: Non-convex set.

(c) The set S is called *balanced*, if it is nonempty and

$$\alpha S \subset S \quad \text{for all } \alpha \in [-1, 1].$$

(d) The set S is called *absolutely convex*, if it is convex and balanced.

Obviously, the empty set is convex and a set which is starshaped at every point is convex as well.

Remark 1.5.

(a) The intersection of arbitrarily many convex sets of a real linear space is convex.

(b) If S and T are nonempty convex subsets of a real linear space X, then the algebraic sum $\alpha S + \beta T$ is convex for all $\alpha, \beta \in \mathbb{R}$. Consequently, for every $\bar{x} \in X$ the translated set $S + \{\bar{x}\}$ is convex as well.

Definition 1.6. Let S be a nonempty subset of a real linear space X. The intersection of all convex subsets of X that contain S is called the *convex hull* of S and is denoted $\mathrm{co}(S)$.

Remark 1.7. For two nonempty subsets S and T of a real linear space we obtain for all $\alpha, \beta \in \mathbb{R}$

$$\mathrm{co}(\alpha S + \beta T) = \alpha \mathrm{co}(S) + \beta \mathrm{co}(T).$$

Next, we consider sets which are algebraically open or closed.

Definition 1.8. Let S be a nonempty subset of a real linear space X.

(a) The set

$$\mathrm{cor}(S) := \{\bar{x} \in S \mid \text{for every } x \in X \text{ there is a } \bar{\lambda} > 0 \text{ with } \bar{x} + \lambda x \in S \ \text{ for all } \lambda \in [0, \bar{\lambda}]\}$$

is called the *algebraic interior* of S (or the *core* of S, see Fig. 1.4).

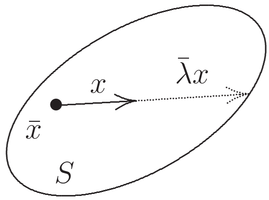

Figure 1.4: $\bar{x} \in \text{cor}(S)$.

(b) The set S with $S = \text{cor}(S)$ is called *algebraically open*.

(c) The set of all elements of X which do not belong to $\text{cor}(S)$ and $\text{cor}(X \backslash S)$ is called the *algebraic boundary* of S.

(d) An element $\bar{x} \in X$ is called *linearly accessible* from S, if there is an $x \in S$, $x \neq \bar{x}$, with the property

$$\lambda x + (1-\lambda)\bar{x} \in S \quad \text{for all } \lambda \in (0,1].$$

The union of S and the set of all linearly accessible elements from S is called the *algebraic closure* of S and it is denoted by

$$\text{lin}(S) := S \cup \{x \in X \mid x \text{ is linearly accessible from } S\}.$$

In the case of $S = \text{lin}(S)$ the set S is called *algebraically closed*.

(e) The set S is called *algebraically bounded*, if for every $\bar{x} \in S$ and every $x \in X$ there is a $\bar{\lambda} > 0$ such that

$$\bar{x} + \lambda x \notin S \quad \text{for all } \lambda \geq \bar{\lambda}.$$

These algebraic notions have a special geometric meaning. Take the intersections of the set S with each straight line in the real linear space X and consider these intersections as subsets of the real line $\mathbb{R}$. Then the set S is algebraically open, if these subsets are open; S is algebraically closed, if these subsets are closed; and S is algebraically bounded, if these subsets are bounded.

Lemma 1.9. *For a nonempty convex subset S of a real linear space we have:*

(a) $\bar{x} \in cor(S),\ \tilde{x} \in lin(S) \implies \{\lambda\tilde{x}+(1-\lambda)\bar{x} | \lambda \in [0,1)\} \subset cor(S)$,

(b) $cor(cor(S)) = cor(S)$,

(c) $cor(S)$ *and* $lin(S)$ *are convex,*

(d) $cor(S) \neq \emptyset \implies lin(cor(S)) = lin(S)$ *and* $cor(lin(S)) = cor(S)$.

A proof of Lemma 1.9 which is rather technical may be found in Kirsch-Warth-Werner [188, p. 9].

Another important class of subsets in a real linear space is introduced in

Definition 1.10. Let C be a nonempty subset of a real linear space X.

(a) The set C is called a *cone*, if

$$x \in C,\ \lambda \geq 0 \implies \lambda x \in C$$

(see Fig. 1.5).

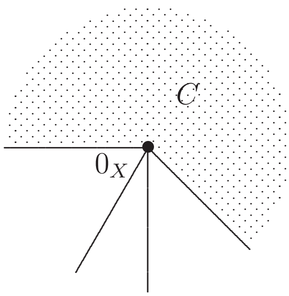

Figure 1.5: Cone.

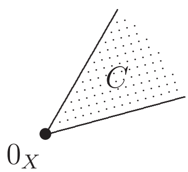

Figure 1.6: Pointed cone.

(b) A cone C is called *pointed*, if

$$C \cap (-C) = \{0_X\}$$

(see Fig. 1.6).

(c) A cone C is called *reproducing*, if

$$C - C = X.$$

In this case one also says that C generates X.

(d) A nonempty convex subset B of a convex cone $C \neq \{0_X\}$ is called a *base* for C, if each $x \in C \backslash \{0_X\}$ has a unique representation of the form

$$x = \lambda b \text{ for some } \lambda > 0 \text{ and some } b \in B$$

(see Fig. 1.7).

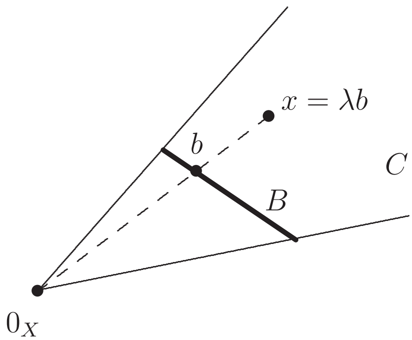

Figure 1.7: Base B for C.

Sometimes a cone is also called a wedge and a pointed wedge is called a cone. But in this book we use the terms in Definition 1.10.

By definition each cone contains the zero element of the real linear space. The simplest cones in a real linear space X are $\{0_X\}$ and X itself. $\{0_X\}$ is also called the *trivial cone*. From a geometric point of view a nontrivial cone is a set of rays emanating from the origin. Consequently, each cone is starshaped at 0_X.

For the investigation of partial orderings convex cones are very important. They are characterized by

Lemma 1.11. *A cone C in a real linear space is convex if and only if*

$$C + C \subset C.$$

Proof.

(a) Assume that C is a convex cone. Then for every $x, y \in C$ we have

$$\frac{1}{2}(x+y) = \frac{1}{2}x + \frac{1}{2}y \in C$$

implying $x + y \in C$. So, the inclusion $C + C \subset C$ is true.

(b) For arbitrary $x, y \in C$ and $\lambda \in [0,1]$ we obtain

$$\lambda x \in C \text{ and } (1-\lambda)y \in C.$$

With the inclusion $C + C \subset C$ we then get

$$\lambda x + (1-\lambda)y \in C,$$

i.e. the cone C is convex.

$\square$

The algebraic interior of a convex cone has interesting properties listed below.

Lemma 1.12. *Let C be a convex cone in a real linear space X with a nonempty algebraic interior. Then:*

(a) $cor(C) \cup \{0_X\}$ is a convex cone,

(b) $cor(C) = C + cor(C)$.

Proof.

(a) Take arbitrary $\bar{x} \in \text{cor}(C)$ and $\mu > 0$. For every $x \in X$ there is a $\bar{\lambda} > 0$ with

$$\bar{x} + \frac{\lambda}{\mu}x \in C \text{ for all } \lambda \in [0, \bar{\lambda}].$$

Since C is a cone, we get

$$\mu\Big(\bar{x} + \frac{\lambda}{\mu} x\Big) = \mu\bar{x} + \lambda x \in C \text{ for all } \lambda \in [0, \bar{\lambda}].$$

So, we obtain $\mu\bar{x} \in \text{cor}(C)$ and with Lemma 1.9, (c) the assertion is obvious.

(b) The inclusion

$$\text{cor}(C) = \{0_X\} + \text{cor}(C) \subset C + \text{cor}(C)$$

is clear. For the proof of the converse inclusion we take arbitrary $\tilde{x} \in C$, $\bar{x} \in \text{cor}(C)$ and $x \in X$. Then there is a $\bar{\lambda} > 0$ with

$$\bar{x} + \lambda x \in C \text{ for all } \lambda \in [0, \bar{\lambda}].$$

Since C is assumed to be convex, we conclude with Lemma 1.11

$$\tilde{x} + \bar{x} + \lambda x \in C \text{ for all } \lambda \in [0, \bar{\lambda}]$$

implying $\tilde{x} + \bar{x} \in \text{cor}(C)$. So, we conclude $C + \text{cor}(C) \subset \text{cor}(C)$.

□

The following lemma gives a sufficient condition for a cone to be reproducing.

Lemma 1.13. *A cone C in a real linear space X is reproducing, if cor$(C) \neq \emptyset$.*

Proof. If $\text{cor}(C)$ is nonempty, take some $\bar{x} \in \text{cor}(C)$ and any $x \in X$. Then there is a $\bar{\lambda} > 0$ with $\bar{x} + \bar{\lambda} x \in C$ implying

$$x \in \frac{1}{\bar{\lambda}} C - \Big\{\frac{1}{\bar{\lambda}}\bar{x}\Big\} \subset C - C.$$

So, we get $X \subset C - C$ and together with the trivial inclusion $C - C \subset X$ we obtain the assertion. □

Next, we turn our attention to the notion of a base B of a nontrivial convex cone. Because of the convexity of B and the uniqueness of λ we have $O_X \notin B$.

Lemma 1.14. *Each nontrivial convex cone with a base in a real linear space is pointed.*

Proof. Let C be a nontrivial convex cone with base B. Take any $x \in C \cap (-C)$ and assume that $x \neq 0_X$. Then there are $b_1, b_2 \in B$ and $\lambda_1, \lambda_2 > 0$ with $x = \lambda_1 b_1 = -\lambda_2 b_2$ implying $\frac{\lambda_1}{\lambda_1+\lambda_2} b_1 + \frac{\lambda_2}{\lambda_1+\lambda_2} b_2 = 0_X \in B$. But this is a contradiction to the afore-mentioned remark. $\Box$

Definition 1.15. Let S be a nonempty subset of a real linear space. The cone

$$\text{cone}(S) := \{x \in X \mid x = \lambda s \text{ for some } \lambda \geq 0 \text{ and some } s \in S\}$$

is called the cone *generated* by S (see Fig. 1.8).

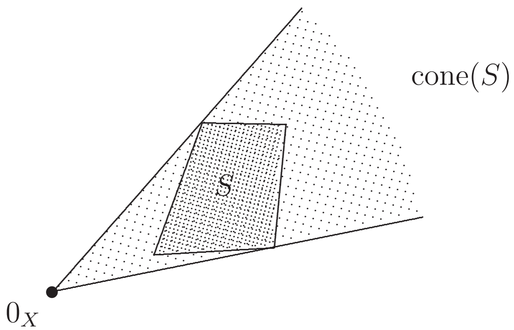

Figure 1.8: Cone generated by S.

It is an important property of a base B of a cone C that $\text{cone}(B) = C$. If $0_X \in \text{cor}(S)$ for a nonempty subset S of a real linear space X, then $\text{cone}(S) = X$.

1.2 Partially Ordered Linear Spaces

In addition to the linear structure of a space we consider a partial ordering which is given in many real linear spaces being of practical interest.

Definition 1.16. Let X be a real linear space.

(a) Each nonempty subset R of the product space $X \times X$ is called a *binary relation* R on X (we write xRy for $(x, y) \in R$).

(b) Every binary relation $\leq$ on X is called a *partial ordering* on X, if the following axioms are satisfied (for arbitrary $w, x, y, z \in X$):

  (i) $x \leq x$;

  (ii) $x \leq y,\ y \leq z \implies x \leq z$;

  (iii) $x \leq y,\ w \leq z \implies x + w \leq y + z$;

  (iv) $x \leq y,\ \alpha \in \mathbb{R}_+ \implies \alpha x \leq \alpha y$.

(c) A partial ordering $\leq$ on X is called *antisymmetric*, if the following implication holds for arbitrary $x, y \in X$:

$$x \leq y,\ y \leq x \implies x = y.$$

In Definition 1.16, (b) with axiom (i) the partial ordering is reflexive and with (ii) it is transitive. The axioms (iii) and (iv) guarantee the compatibility of the partial ordering with the linear structure of the space.

Definition 1.17. A real linear space equipped with a partial ordering is called a *partially ordered linear space.*

It is important to note that in a partially ordered linear space two arbitrary elements cannot be compared, in general, in terms of the partial ordering. A significant characterization of a partial ordering in a real linear space is given by

Theorem 1.18. *Let X be a real linear space.*

(a) If $\leq$ is a partial ordering on X, then the set

$$C := \{x \in X \mid 0_X \leq x\}$$

is a convex cone. If, in addition, $\leq$ is antisymmetric, then C is pointed.

(b) If C is a convex cone in X, then the binary relation

$$\leq_C := \{(x, y) \in X \times X \mid y - x \in C\}$$

is a partial ordering on X. If, in addition, C is pointed, then $\leq_C$ is antisymmetric.

This theorem is easy to prove and is of great importance because a partial ordering can be investigated using convex analysis.

The next definition is based on the result of Theorem 1.18.

Definition 1.19. A convex cone characterizing a partial ordering in a real linear space is called an *ordering cone*.

Several authors also call an ordering cone a *positive cone*. We denote $\leq_C$ as a partial ordering induced by a convex cone C.

Example 1.20. For $X = \mathbb{R}^n$ the ordering cone of the componentwise partial ordering on $\mathbb{R}^n$ is given by

$$C := \{x \in \mathbb{R}^n \mid x_i \geq 0 \text{ for all } i \in \{1, \ldots, n\}\} = \mathbb{R}^n_+.$$

It is also called the natural ordering cone. Other ordering cones in $\mathbb{R}^n$ are for instance

$$\{x \in \mathbb{R}^n \mid x_i \geq 0 \text{ for all } i \in \{1, \ldots, m\} \text{ and } x_i = 0 \text{ for all } i \in \{m+1, \ldots, n\}\} \quad \text{for some } 1 \leq m < n$$

or $\{0_{\mathbb{R}_n}\}$ and $\mathbb{R}^n$ itself. $\mathbb{R}_+$, $\mathbb{R}_-$, $\{0\}$ and $\mathbb{R}$ are the only ordering cones in $\mathbb{R}$. Ordering cones of special infinite dimensional linear spaces will be presented in Subsection 1.4.

Definition 1.21. Let X be a partially ordered linear space. For arbitrary elements $x, y \in X$ with $x \leq y$ the set

$$[x, y] := \{z \in X \mid x \leq z \leq y\}$$

is called the *order interval* between x and y (see Fig. 1.9)