

GENERAL THEOREMS OF THE ALTERNATIVE IN VARIABLE ORDERING STRUCTURES

JOHANNES JAHN

Department Mathematik, Universität Erlangen-Nürnberg, Cauerstr. 11, 91058 Erlangen, Germany

Abstract. This paper presents two new theorems of the alternative, which can be used for the proof of necessary optimality conditions for nonconvex optimization problems like discrete-continuous problems and for vector optimization problems with variable orderings.

Keywords. Nonconvex optimization; Farkas' lemma; Necessary optimality conditions; Theorems of the alternative; Vector optimization.

1. INTRODUCTION

In 1902, Farkas [1] published a famous result, which is nowadays called Farkas' lemma. This lemma concerns linear inequalities and it is formulated in the form of an alternative, i.e. exactly one of two statements is true. Among others, the Farkas' lemma can be used for the proof of the well-known Karush-Kuhn-Tucker conditions in nonlinear optimization.

During the last decades, the Farkas' lemma has been extended and generalized to further theorems of the alternative. First results were given by Lehmann and Oettli [2] in finite dimensions. Giannessi [3] proved a version with a set-valued map defined on a Banach space. An overview of further developments was given by Dinh and Jeyakumar [4]. In general, the investigations in [4] work with generalized convexity assumptions.

This paper goes in a different direction because it presents two general theorems of the alternative without any generalized convexity assumption on the considered map. The preimage of this map may also be nonconvex. The only convexity assumption concerns the image sets of a set-valued map defining a variable ordering structure. Since this ordering structure is not fixed but variable, these new theorems of the alternative have cutting-edge applications.

If one works with variable ordering structures, one cannot simply extend the results for fixed orderings to this more general case. This leads to the fact that separation theorems are used pointwise, i.e., the separating functionals depend on the used variable. If one applies these results to necessary optimality conditions, one generally gets more than one Lagrange multiplier per constraint. This is a well-known fact in discrete-continuous nonlinear optimization (see, e.g., [5]).

This short paper is organized as follows: The two new general theorems of the alternative are given in Section 2. Applications of these results, like the classical Farkas' lemma and necessary

E-mail address: johannes.jahn@fau.de.

Received June 12, 2020; Accepted November 1, 2020.

optimality conditions in discrete-continuous nonlinear optimization and in vector optimization with variable orderings, are presented in Section 3.

2. MAIN RESULTS

Two variants of a general theorem of the alternative are presented in this section. For the first one, we use the following assumption.

Assumption 2.1. Let S be a nonempty subset of a real linear space X , let Y be a locally convex topological real linear space, let $C : S \rightrightarrows Y$ be a set-valued map with image sets being closed convex cones, and let $f : S \rightarrow Y$ be a single-valued map.

Under this assumption, we present the first general theorem of the alternative.

Theorem 2.1. *Under Assumption 2.1, exactly one of the two statements holds:*

- (A) $\exists x \in S : f(x) \in -C(x)$
- (B) $\forall x \in S \exists \ell^x \in C^*(x) \setminus \{0_{Y^*}\} : \ell^x(f(x)) > 0$
($C^(x)$ denotes the dual cone of $C(x)$ for all $x \in S$).*

Proof.

(a) Assume that statement (A) is true, i.e. there is some $x \in S$ with $f(x) \in -C(x)$. Then, we conclude

$$\ell^x(f(x)) \leq 0 \text{ for all } \ell^x \in C^*(x).$$

Consequently, the statement (B) does not hold.

(b) Now, assume that statement (A) is not true. Then we have

$$-f(x) \notin C(x) \text{ for all } x \in S.$$

Let $x \in S$ be arbitrarily chosen. By a strict separation theorem (compare [6, Thm. 3.18]), there is a continuous linear functional $\ell^x \in Y^* \setminus \{0_{Y^*}\}$ and a real number α^x with

$$\ell^x(-f(x)) < \alpha^x \leq \ell^x(c^x) \text{ for all } c^x \in C(x). \quad (2.1)$$

If we assume that $\ell^x \notin C^*(x) \setminus \{0_{Y^*}\}$, then there exists some $c^x \in C(x)$ with $\ell^x(c^x) < 0$, which contradicts the right inequality in (2.1) because $C(x)$ is a cone. So, we have

$$\ell^x \in C^*(x) \setminus \{0_{Y^*}\}.$$

For $c^x = 0_Y$, we find that (2.1) implies $\ell^x(f(x)) > 0$. Consequently, the statement (B) is true. \square

For another variant of Theorem 2.1, we slightly change Assumption 2.1.

Assumption 2.2. Let S be a nonempty subset of a real linear space X , let Y be a topological real linear space, let $C : S \rightrightarrows Y$ be a set-valued map with image sets being convex cones with nonempty interior, and let $f : S \rightarrow Y$ be a single-valued map.

Here we change an assumption of the variable order structure. Now we do not require that the convex cones $C(x)$ for arbitrary $x \in S$ are closed but we assume that the interior $\text{int}(C(x))$ is nonempty.

Theorem 2.2. *Under Assumption 2.2, exactly one of the two statements holds:*

- (C) $\exists x \in S : f(x) \in -\text{int}(C(x))$

$$(D) \quad \forall x \in S \exists \ell^x \in C^*(x) \setminus \{0_{Y^*}\} : \ell^x(f(x)) \geq 0.$$

Proof.

(a) Assume that statement (C) is true, i.e. there is some $x \in S$ with $f(x) \in -\text{int}(C(x))$. By [6, Lemma 3.21,(c)], we have

$$\text{int}(C(x)) = \{y \in Y \mid \ell^x(y) > 0 \text{ for all } \ell^x \in C^*(x) \setminus \{0_{Y^*}\}\}.$$

Hence, we conclude

$$\ell^x(f(x)) < 0 \text{ for all } \ell^x \in C^*(x) \setminus \{0_{Y^*}\}.$$

So, statement (D) is not true.

(b) Now assume that statement (C) is not true. Then it follows

$$\{-f(x)\} \cap \text{int}(C(x)) = \emptyset \text{ for all } x \in S.$$

Choose an arbitrary $x \in S$. By Eidelheit's separation theorem (compare [6, Thm. 3.16]), there are a continuous linear functional $\ell^x \in Y^* \setminus \{0_{Y^*}\}$ and a real number α^x with

$$\ell^x(c^x) \geq \alpha^x \geq \ell^x(-f(x)) \text{ for all } c^x \in C(x). \quad (2.2)$$

Since the set $C(x)$ is a cone, we get $\ell^x \in C^* \setminus \{0_{Y^*}\}$. For $c^x = 0_Y$, we have that (2.2) implies $\ell^x(f(x)) \geq 0$. Consequently, statement (D) is true. \square

3. THE APPLICATIONS TO OPTIMIZATION

In this section, we apply the two presented theorems of the alternative to different issues of optimization. We derive the classical Farkas' lemma and a necessary optimality condition in discrete-continuous optimization as well as in vector optimization with variable ordering structures.

3.1. Farkas' lemma. A direct consequence of Theorem 2.1 is the well-known Farkas' lemma in optimization.

Corollary 3.1 (Farkas' lemma). *For an arbitrary matrix $A \in \mathbb{R}^{(m,n)}$ (with $m, n \in \mathbb{N}$) and an arbitrary vector $b \in \mathbb{R}^m$, exactly one of the two statements holds:*

$$(E) \quad \exists x \geq 0_{\mathbb{R}^n} : Ax = b$$

$$(F) \quad \exists y \in \mathbb{R}^m : A^T y \geq 0_{\mathbb{R}^n} \text{ and } b^T y < 0$$

(\geq has to be understood in a componentwise sense).

Proof. In Theorem 2.1 we set $X := S := \mathbb{R}^n$, $Y := \mathbb{R}^m$, $C(x) := K := \{A\tilde{x} \mid \tilde{x} \geq 0_{\mathbb{R}^n}\}$ for all $x \in \mathbb{R}^n$ and $f(x) = -b$ for all $x \in \mathbb{R}^n$. Then the statement (E) is equivalent to $b \in K$, which is again equivalent to the statement (A) in Theorem 2.1 because the maps f and C are constant. In this special case, the statement (B) in Theorem 2.1 can be written as

$$\forall x \in \mathbb{R}^n \exists y \in K^* \setminus \{0_{\mathbb{R}^m}\} : y^T(-b) > 0.$$

So, there is some $y \in K^*$ with $y^T b < 0$. This statement is equivalent to the statement (F) because the dual cone of K is given as

$$K^* = \{y \in \mathbb{R}^m \mid A^T y \geq 0_{\mathbb{R}^n}\},$$

which completes the proof. \square

Using a strict separation theorem the proof of the general Theorem 2.1 is simpler than the standard proof of the Farkas' lemma generally given in linear programming (compare [7, Thm. 2.2.3] among others). It is well-known that several other theorems of the alternative can be derived from the Farkas' lemma (see [7, p. 17–19]).

3.2. Necessary optimality conditions in discrete-continuous nonlinear optimization. Now we apply Theorem 2.1 to the problems of nonconvex optimization, which also include discrete-continuous nonlinear problems. The standard assumption of this subsection is as follows.

Assumption 3.1. Let \hat{S} be a nonempty subset of a real linear space X , let Y and Z be locally convex topological real linear spaces, let C be a closed convex cone in Y , let $f : \hat{S} \rightarrow \mathbb{R}$ be a given functional, let $g : \hat{S} \rightarrow Y$ and $h : \hat{S} \rightarrow Z$ be given maps, and let the constraint set

$$S := \{x \in \hat{S} \mid g(x) \in -C, \ h(x) = 0_Z\}$$

be nonempty.

Under this assumption, we consider the nonlinear optimization problem with inequality and equality constraints

$$\begin{aligned} & \min f(x) \\ & \text{subject to} \\ & g(x) \in -C, \\ & h(x) = 0_Z, \\ & x \in \hat{S}. \end{aligned} \tag{3.1}$$

Since the superset \hat{S} is not assumed to be convex, this problem class also subsumes nonlinear discrete-continuous optimization problems. The discrete variables can be defined by an appropriate choice of the set \hat{S} (for instance, $\hat{S} := \mathbb{Z}^m \times \mathbb{R}^n$ for $m, n \in \mathbb{N}$ in mixed-integer programming).

Corollary 3.2. *Let Assumption 3.1 be satisfied. If $\bar{x} \in S$ is a minimal solution of the optimization problem (3.1), then, for every $x \in \hat{S}$ with $f(x) \neq f(\bar{x})$, there are continuous linear functionals $\ell_g^x \in C^*$ and $\ell_h^x \in Z^*$ with*

$$f(x) + \ell_g^x(g(x)) + \ell_h^x(h(x)) > f(\bar{x})$$

(C^* denotes the dual cone of C).

Proof. Let $\bar{x} \in S$ be an arbitrary minimal solution of problem (3.1), i.e. there is no $x \in S$ with $f(x) < f(\bar{x})$. Then there is no $x \in \hat{S}$, $f(x) \neq f(\bar{x})$, with

$$\begin{pmatrix} f(x) - f(\bar{x}) \\ g(x) \\ h(x) \end{pmatrix} \in (-\mathbb{R}_+) \times (-C) \times \{0_Z\}$$

(\mathbb{R}_+ denotes the set of all non-negative real numbers). The product cone $\mathbb{R}_+ \times C \times \{0_Z\}$ is closed and convex and its dual cone can be written as $\mathbb{R}_+ \times C^* \times Z^*$. If we apply Theorem 2.1, then the statement (B) is true, i.e. for every $x \in \hat{S}$ with $f(x) \neq f(\bar{x})$, there are a real number $\mu^x \geq 0$ and continuous linear functionals $\ell_g^x \in C^*$ and $\ell_h^x \in Z^*$ with $(\mu^x, \ell_g^x, \ell_h^x) \neq (0, 0_{Y^*}, 0_{Z^*})$ and

$$\mu^x(f(x) - f(\bar{x})) + \ell_g^x(g(x)) + \ell_h^x(h(x)) > 0. \tag{3.2}$$

If $\mu^x > 0$, then

$$f(x) + \frac{1}{\mu^x} \ell_g^x(g(x)) + \frac{1}{\mu^x} \ell_h^x(h(x)) > f(\bar{x})$$

and with $\frac{1}{\mu^x} \ell_g^x \in C^*$, the assertion is shown. Otherwise, if $\mu^x = 0$, then

$$\ell_g^x(g(x)) + \ell_h^x(h(x)) > 0$$

follows from (3.2). In this case, there is some $\bar{\mu}^x > 0$ so that

$$\bar{\mu}^x (f(x) - f(\bar{x})) > \underbrace{-\ell_g^x(g(x)) - \ell_h^x(h(x))}_{<0},$$

which implies

$$f(x) + \frac{1}{\bar{\mu}^x} \ell_g^x(g(x)) + \frac{1}{\bar{\mu}^x} \ell_h^x(h(x)) > f(\bar{x}).$$

This completes the proof. \square

Remark 3.1. The necessary optimality condition in Corollary 3.2 holds under weak assumptions. The functional f and the maps g and h do not need to be convex. Moreover the superset \hat{S} is not assumed to be convex. And a constraint qualification is not required. On the other hand the “Lagrange multipliers” depend on $x \in \hat{S}$ with $f(x) \neq f(\bar{x})$, which makes the approach more complicated in practice. It is well-known in discrete-continuous nonlinear optimization that one generally has more than one multiplier per constraint (compare [5]).

Corollary 3.2 extends [8, Thm. 3.2] to equality constraints. Its proof partially uses an approach in the proof of [9, Satz 6.11], which is based on [10] under generalized convexity assumptions.

3.3. Vector optimization with variable orderings. Vector optimization problems with variable ordering structures form a large class of optimization problems, which also covers problems where objectives and constraints depend on certain conditions. If such a condition is not fulfilled, then the corresponding objective or constraint vanishes. Such optimization problems occur in structural optimization [11] and image registration of medical data ([12], [13, Subsection 1.3.1]).

Assumption 3.2. Let \hat{S} be a nonempty subset of a real linear space X , let Y and Z be topological real linear spaces, let $f : \hat{S} \rightarrow Z$ and $g : \hat{S} \rightarrow Y$ be given vector functions, let $C_f : \hat{S} \rightrightarrows Z$ and $C_g : \hat{S} \rightrightarrows Y$ be set-valued maps with image sets being convex cones with nonempty interior, and let the constraint set

$$S := \{x \in \hat{S} \mid g(x) \in -C_g(x)\}$$

be nonempty.

The set-valued maps C_f and C_g define a variable ordering structure in the image spaces Z and Y , respectively. Based on these maps, Eichfelder [13] introduced different types of variable ordering structures. Here we only use the order relations \leq_f and \leq_g defined by (for arbitrary $x_1, x_2 \in \hat{S}$)

$$f(x_1) \leq_f f(x_2) :\iff f(x_2) - f(x_1) \in C_f(x_1)$$

and

$$g(x_1) \leq_g g(x_2) :\iff g(x_2) - g(x_1) \in C_g(x_1).$$

In this case, the inequality constraint $g(x) \in -C_g(x)$ for arbitrary $x \in \hat{S}$ in Assumption 3.2 is equivalent to the inequality constraint $g(x) \leq_g 0_Y$.

Under Assumption 3.2, we consider the vector optimization problem with variable orderings

$$\begin{aligned} & \min_{C_f(x)} f(x) \\ & \text{subject to} \\ & g(x) \in -C_g(x), \\ & x \in \hat{S}. \end{aligned} \tag{3.3}$$

Definition 3.1. Let Assumption 3.2 be satisfied. A feasible element $\bar{x} \in S$ is called a *weak minimal solution* of the vector optimization problem with variable orderings (3.3) if the image $f(\bar{x})$ is a weak nondominated element of the image set $f(S)$, i.e.,

$$\nexists x \in S: f(\bar{x}) \in \{f(x)\} + \text{int}(C_f(x)). \tag{3.4}$$

The condition (3.4) can be interpreted in such a way that there is no feasible element $x \in S$ so that $f(x)$ is “strictly less” than $f(\bar{x})$.

Corollary 3.3. Let Assumption 3.2 be satisfied. If $\bar{x} \in S$ is a weak minimal solution of the vector optimization problem (3.3), then, for every $x \in \hat{S}$, there are continuous linear functionals $\ell_f^x \in C_f^*(x)$ and $\ell_g^x \in C_g^*(x)$ with $(\ell_f^x, \ell_g^x) \neq (0_{Z^*}, 0_{Y^*})$ and

$$\ell_f^x(f(x) - f(\bar{x})) + \ell_g^x(g(x)) \geq 0.$$

Proof. Let $\bar{x} \in S$ be an arbitrary weak minimal solution of the vector optimization problem (3.3), i.e.,

$$\nexists x \in S: f(x) - f(\bar{x}) \in -\text{int}(C_f(x)).$$

Then there is no $x \in \hat{S}$ such that

$$\begin{pmatrix} f(x) - f(\bar{x}) \\ g(x) \end{pmatrix} \in [-\text{int}(C_f(x))] \times [-C_g(x)].$$

Consequently, there is no $x \in \hat{S}$ such that

$$\begin{pmatrix} f(x) - f(\bar{x}) \\ g(x) \end{pmatrix} \in [-\text{int}(C_f(x))] \times [-\text{int}(C_g(x))].$$

Next, we apply Theorem 2.2 and the statement (D) holds. In this special case, this means that, for every $x \in \hat{S}$, there are continuous linear functionals $\ell_f^x \in C_f^*(x)$ and $\ell_g^x \in C_g^*(x)$ with $(\ell_f^x, \ell_g^x) \neq (0_{Z^*}, 0_{Y^*})$ and

$$\ell_f^x(f(x) - f(\bar{x})) + \ell_g^x(g(x)) \geq 0,$$

which has to be shown. □

The proof of Corollary 3.3 partially uses an approach given in the proof of [9, Satz 6.11] (based on [10]), which is done under generalized convexity assumptions. Notice that a similar necessary optimality condition can be shown using Theorem 2.1.

4. CONCLUSION

Although the theorems of the alternative are known for nearly 120 years, this paper presents new general versions, which can be used in very modern fields of optimization, like discrete-continuous problems or problems with variable ordering structures. Hence, general theorems of the alternative are cutting-edge results because they can be used to derive important necessary optimality conditions with more than one Lagrange multiplier per constraint.

REFERENCES

- [1] J. Farkas, Theorie der einfachen Ungleichungen, J. Reine Angew. Math. 124 (1902), 1-27.
- [2] R. Lehmann, W. Oettli, The theorem of the alternative, the key theorem, and the vector-maximum problem, Math. Program. 8 (1975), 332-344.
- [3] F. Giannessi, Theorems of the alternative for multifunctions with applications to optimization: General results, J. Optim. Theory Appl. 55 (1987), 233-256.
- [4] N. Dinh, V. Jeyakumar, Farkas' lemma: three decades of generalizations for mathematical optimization, Top 22 (2014), 1-22.
- [5] J. Jahn, M. Knossalla, Lagrange theory of discrete-continuous nonlinear optimization, J. Nonlinear Var. Anal. 2 (2018), 317-342.
- [6] J. Jahn, Vector Optimization - Theory, Applications, and Extensions, Springer, Berlin, 2011.
- [7] D.T. Luc, Multiobjective Linear Programming - An Introduction, Springer, Cham, 2016.
- [8] J. Jahn, A.A. Khan, Optimality conditions in nonlinear vector optimization with variable ordering structures, Pure Appl. Funct. Anal. to appear.
- [9] D. Stadter, Optimalitätskriterien in der Vektroptimierung, master thesis, University of Erlangen-Nürnberg, Erlangen, Germany, 2016.
- [10] X.M. Yang, Alternative theorems and optimality conditions with weakened convexity, Opsearch, 29 (1992), 115-125.
- [11] W. Achtziger, C. Kanzow, Mathematical programs with vanishing constraints: optimality conditions and constraint qualifications, Math. Program. Ser. A 114 (2008), 69-99.
- [12] M. Wacker, Multikriterielle Optimierung bei Registrierung medizinischer Daten, master thesis, University of Erlangen-Nürnberg, Erlangen, Germany, 2008.
- [13] G. Eichfelder, Variable Ordering Structures in Vector Optimization, Springer, Berlin, 2014.