Minimax Theory and its Applications
Volume 6 (2021), No. 1, 127–144

Optimality Conditions in Discrete-Continuous Nonlinear Optimization

Gabriele Eichfelder
Institut für Mathematik, TU Ilmenau, 98693 Ilmenau, Germany
gabriele.eichfelder@tu-ilmenau.de

Johannes Jahn
Department Mathematik, Universität Erlangen-Nürnberg, 91058 Erlangen, Germany
johannes.jahn@fau.de

Received: May 5, 2020
Accepted: July 27, 2020

We present necessary and sufficient optimality conditions for discrete-continuous nonlinear optimization problems including mixed-integer nonlinear problems. This theory does not utilize an extension of the Lagrange theory of continuous optimization but it works with certain max functionals for a separation of two sets where one of them is nonconvex. These functionals have the advantage that they can be used for nonconvex optimization problems. This theory avoids getting several Lagrange multipliers per constraint.

Keywords: Nonlinear optimization, optimality conditions, discrete-continuous variables, mixed-integer nonlinear problems.

2010 Mathematics Subject Classification: 90C30, 90C46, 90C11.

1. Introduction

In discrete-continuous nonlinear optimization it is well-known that one generally works with several Lagrange multipliers per constraint (e.g. compare [1],[9]). This fact is based on the Lagrange theory, which is extended to discrete-continuous problems (see [9]). It is interesting to note that there are examples with only a small number of Lagrange multipliers per constraint but there may also be problems with many multipliers per constraint. The higher the number of Lagrange multipliers per constraint the more complex are the optimality conditions.

The present paper develops a completely different approach to optimality conditions. In contrast to the standard theory we characterize a unique minimal solution of a discrete-continuous optimization problem. Although this seems to be restrictive, notice that Planiden and Wang [12] have shown that the most convex functions have unique minimizers.

Instead of linear separation of sets the present theory works with a specific nonconvex separation of sets. If one describes the boundary of the negative ordering cone used for the definition of the inequality constraints by an appropriate functional, then one obtains an optimality condition with a certain max functional. This max functional

ISSN 2199-1413 (printed), ISSN 2199-1421 (electronic) / $ 2.50

plays the role of a separating functional where one of the two separated sets may be nonconvex.

The description of the boundary of a negative ordering cone by a certain functional is well-known in vector optimization and it is closely related to scalarization results and nonconvex separation (e.g. see [7], [5] and [13], among others). These investigations on scalarization in vector optimization are very helpful for the formulation of optimality conditions in discrete-continuous optimization.

Since discrete-continuous optimization problems have a nonconvex constraint set, nonconvex separation seems to be an essential tool for the formulation of optimality conditions. If the number of functionals defining the inequality constraints is finite and the ordering cone is the positive orthant, then the resulting separating functional has a simple form. In this special case a multiplier rule can be given as a sufficient optimality condition. A multiplier rule as a necessary optimality condition can only be formulated under strong assumptions. But these optimality conditions have the advantage that one has only one multiplier per constraint.

This paper is organized as follows: After this introduction Section 2 presents basics and an introducing example. Optimality conditions for a unique minimal solution are given in the third section. This section starts with quite general multiplier-free conditions, which are finally specialized to multiplier rules.

2. Basic remarks and example

The standard notions used in this paper are as follows: The algebraic dual space of a real linear space X is termed X' whereas X^* describes the topological dual space of a real linear topological space X. $\dim(\cdot)$ denotes the dimension of a real linear space. The set of all positive real numbers is called $\mathbb{R}_{++}$ and the set of all nonnegative real numbers is termed $\mathbb{R}_+$. $\mathrm{epi}(\cdot)$, $\mathrm{conv}(\cdot)$, $\mathrm{bd}(\cdot)$, $\mathrm{int}(\cdot)$ and $\mathrm{cor}(\cdot)$ abbreviate the epigraph, the convex hull, the boundary, the topological interior and the algebraic interior (or core) of a set, respectively. For some element $\bar{x}$ of a subset A of a real linear space X the normal cone to A at $\bar{x}$ is denoted by

$$N(A,\bar{x}) := \{\ell \in X' \mid \ell(x-\bar{x}) \leq 0 \text{ for all } x \in A\}.$$

Throughout this paper a nonempty subset S_d of a real linear space X_d is said to be *discrete*, iff $S_d = \{x_d^i\}_{i\in N}$ for $N := \{1,2,\dots,n\}$ (with $n \in \mathbb{N}$) or $N := \mathbb{N}$ with elements $x_d^1, x_d^2, \dots \in X_d$.

Assumption 2.1. Let X_d, X_c, Y and Z be real linear spaces, and let $C \subset Y$ be a convex cone in Y. Let S_d be a nonempty discrete subset of X_d, and let S_c be a nonempty subset of X_c. Let $f : S_d \times S_c \to \mathbb{R}$, $g : S_d \times S_c \to Y$ and $h : S_d \times S_c \to Z$ be given maps and let the constraint set

$$S := \{(x_d, x_c) \in S_d \times S_c \mid g(x_d, x_c) \in -C,\ h(x_d, x_c) = 0_Z\}$$

be nonempty.

For simplicity the subscripts d and c clarify the discrete-continuous structure, i.e. x_d and x_c are the discrete and continuous variables belonging to the sets S_d and

S_c, respectively, and these sets are subsets of the real linear spaces X_d and X_c, respectively.

Under Assumption 2.1 we investigate the discrete-continuous optimization problem

$$\min_{(x_d,\,x_c)\in S} f(x_d, x_c). \tag{1}$$

In [1] optimality conditions are shown in mixed-integer optimization where one obtains several multipliers per constraint (in contrast to only one multiplier per constraint in continuous optimization). Based on a theorem given by Doignon [4] it is stated in [1, p. 558] that the term $2^{\dim(X_d)}$ is an upper bound on the number of multipliers per constraint in mixed-integer optimization.

An extension of the known Lagrange theory to discrete-continuous optimization given in [9] also leads to optimality conditions with several multipliers per constraint. We recall the known necessary optimality condition for the optimization problem (1).

Theorem 2.2. [9, Thm. 3.1] *Let Assumption* 2.1 *be satisfied and in addition, let* X_c, Y *and* Z *be real topological linear spaces with* $\mathrm{int}(S_c) \neq \emptyset$ *and* $\mathrm{int}(C) \neq \emptyset$. *Let* $\bar{x} = (x_d^j, \bar{x}_c)$ *(for some* $j \in N$*) be a minimal solution of the discrete-continuous optimization problem* (1). *For every* $i \in N$ *let the set*

$$B_i := \left\{ \begin{pmatrix} f(x_d^i, x_c) - f(\bar{x}) + \alpha \\ g(x_d^i, x_c) + y \\ h(x_d^i, x_c) \end{pmatrix} \in \mathbb{R}\times Y\times Z \;\middle|\; x_c \in \mathrm{int}(S_c),\ \alpha > 0,\ y \in \mathrm{int}(C) \right\} \tag{2}$$

be convex and let $h(x_d^i, \mathrm{int}(S_c))$ *be an open set. Then for every* $i \in N$ *there exist a real number* $\mu^i \geq 0$ *and continuous linear functionals* $\ell_g^i \in C^*$ *and* $\ell_h^i \in Z^*$ *with* $(\mu^i, \ell_g^i, \ell_h^i) \neq (0, 0_{Y^*}, 0_{Z^*})$, *and the inequality*

$$0 \leq \inf_{i\in N} \left\{ \mu^i \left(f(x_d^i, x_c) - f(\bar{x}) \right) + \ell_g^i \left(g(x_d^i, x_c) \right) + \ell_h^i \left(h(x_d^i, x_c) \right) \right\} \text{ for all } x_c \in S_c \tag{3}$$

and the equality

$$\ell_g^j \left(g(\bar{x}) \right) = 0 \tag{4}$$

are fulfilled.

Until now it is unknown how many continuous linear functionals ℓ_g^i and ℓ_h^i we really need for the multiplier rule (3), (4). The following example illustrates this point.

Example 2.3. We investigate the discrete-continuous optimization problem

$$\begin{array}{rl} \min & \frac{1}{2}(x_d^i-2)(x_d^i-3)(x_d^i-4) - \frac{1}{2}(x_d^i-1)(x_d^i-3)(x_d^i-4) \\ & +2(x_d^i-1)(x_d^i-2)(x_d^i-4) - \frac{3}{2}(x_d^i-1)(x_d^i-2)(x_d^i-3) + x_c \\ & \text{subject to} \\ & x_d^i - x_c - 3 \leq 0, \quad (x_d^i, x_c) \in \underbrace{\{1,\ldots,4\}}_{=:S_d} \times \underbrace{\left[-\frac{1}{2}, \frac{1}{2}\right]}_{=:S_c}. \end{array} \tag{5}$$

It can be easily checked that the point $\bar{x} := (3, 0)$ is the minimal solution of this problem with the minimal value $f(\bar{x}) = -4$. For every $i \in \{1, \ldots, 4\}$ the set B_i in (2) can be written as

$$B_i = \left\{ \left(\begin{array}{c} f(x_d^i, x_c) + 4) \\ x_d^i - x_c - 3 \end{array} \right) \in \mathbb{R}^2 \;\middle|\; x_c \in (-\tfrac{1}{2}, \tfrac{1}{2}) \right\} + \mathbb{R}_{++} \times \mathbb{R}_{++}.$$

Then we obtain

$$B_1 = \left\{ \begin{pmatrix} 1 \\ -2 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \end{pmatrix} x_c \;\middle|\; x_c \in (-\tfrac{1}{2}, \tfrac{1}{2}) \right\} + \mathbb{R}^2_{++},$$

$$B_2 = \left\{ \begin{pmatrix} 3 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \end{pmatrix} x_c \;\middle|\; x_c \in (-\tfrac{1}{2}, \tfrac{1}{2}) \right\} + \mathbb{R}^2_{++},$$

$$B_3 = \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} x_c \;\middle|\; x_c \in (-\tfrac{1}{2}, \tfrac{1}{2}) \right\} + \mathbb{R}^2_{++}$$

and
$$B_4 = \left\{ \begin{pmatrix} -5 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \end{pmatrix} x_c \;\middle|\; x_c \in (-\tfrac{1}{2}, \tfrac{1}{2}) \right\} + \mathbb{R}^2_{++}.$$

Figure 2.1 illustrates these four sets. It is obvious from Figure 2.1 that H_2 is the

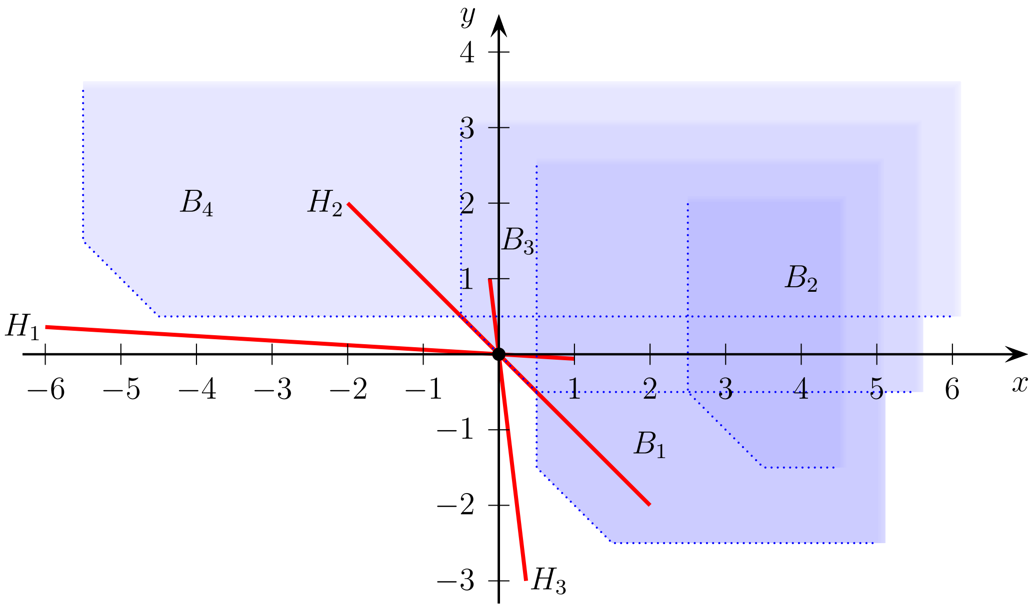

Figure 2.1: Illustration of the sets $B_1, \ldots, B_4$ and the separating hyperplanes H_1, H_2, H_3 in Example 2.3.

unique hyperplane separating the sets $\{0_{\mathbb{R}^2}\}$ and B_3. The hyperplane H_1 separates the sets $\{0_{\mathbb{R}^2}\}$ and B_4 whereas the hyperplane H_3 separates the sets $\{0_{\mathbb{R}^2}\}$ and B_1 and it separates the sets $\{0_{\mathbb{R}^2}\}$ and B_2. Hence, we need at least three hyperplanes, i.e. in Theorem 2.2 we have to work with at least three separating linear functionals (compare the proof in [9, Thm. 3.1]).

Remark 2.4.

(a) Example 2.3 shows that one needs at least three multipliers per constraint although the integer variables are one-dimensional (i.e. $\dim(X_d) = 1$). So, for problems with a constraint set S given in Assumption 2.1, the known upper bound $2^{\dim(X_d)}$ in mixed-integer optimization (as stated in [1, p. 558]) is not transferable to this more general case. Notice that Lagrange multipliers in Theorem 2.2 are obtained by separation in the image product space of the objective functional and the constraint maps and not in a preimage space.

(b) It is evident in Figure 2.1 that a separation of the zero element and the nonconvex set $\cup_{i=1}^{4} B_i$ can be simpler done by a nonlinear functional. If the hyperplanes H_1 and H_3 are used for the definition of an appropriate nonlinear separating functional ξ, then we obtain the requested separation. This point is illustrated in Figure 2.2 and it motivates the investigation in the next section.

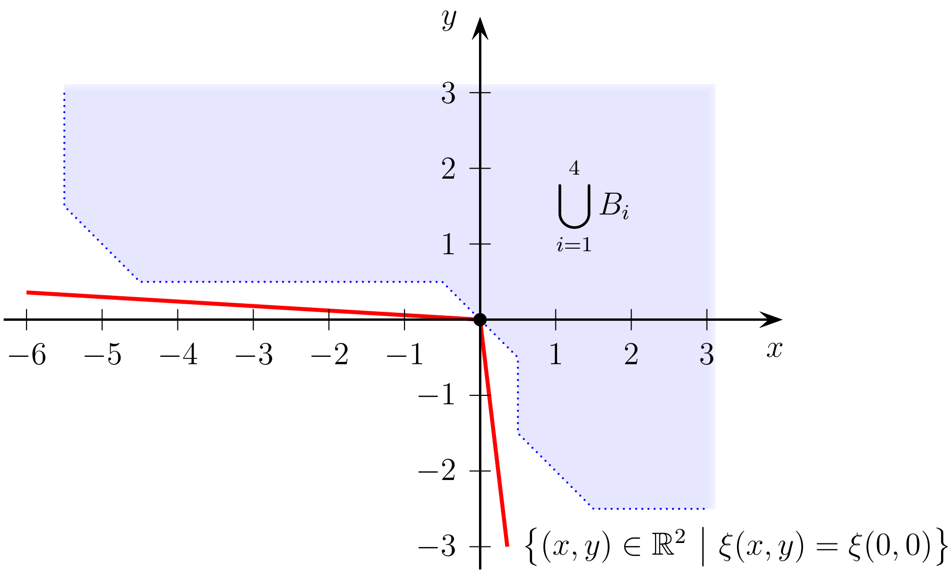

Figure 2.2: Illustration of a nonlinear separation as indicated in Remark 2.4(b).

3. Optimality conditions for a unique minimal solution

Example 2.3 shows that a Lagrange approach in discrete-continuous optimization can lead to a large number of multipliers per constraint. This example also motivates the following multiplier-free approach, which uses a different type of separation.

The approach of this section is developed for unique minimal solutions. Recall that $\bar{x} \in S$ is said to be a *unique minimal solution* of the optimization problem (1) iff

$$f(\bar{x}) < f(x) \text{ for all } x \in S \text{ with } x \neq \bar{x}.$$

We start with a simple basic optimality condition.

Lemma 3.1. *Let Assumption* 2.1 *be satisfied, and let some* $\bar{x} \in S$ *be given.*

(a) *The vector* $\bar{x}$ *is a unique minimal solution of problem (1) if and only if*

$$\big[-\mathbb{R}_+ \times (-C) \times \{0_Z\}\big] \cap \left\{ \begin{pmatrix} f(x) - f(\bar{x}) \\ g(x) \\ h(x) \end{pmatrix} \in \mathbb{R} \times Y \times Z \;\middle|\; x \in S_d \times S_c,\ x \neq \bar{x} \right\} = \emptyset. \quad (6)$$

(b) *The vector* $\bar{x}$ *is a minimal solution of problem (1) if and only if*

$$\big[-\mathbb{R}_{++} \times (-C) \times \{0_Z\}\big] \cap \left\{ \begin{pmatrix} f(x) - f(\bar{x}) \\ g(x) \\ h(x) \end{pmatrix} \in \mathbb{R} \times Y \times Z \;\middle|\; x \in S_d \times S_c \right\} = \emptyset.$$

Proof. (a)(i) Let $\bar{x}$ be a unique minimal solution of the discrete-continuous optimization problem (1). Then we have for all $x \in S_d \times S_c$ with $x \in S$ and $x \neq \bar{x}$

$$f(x) - f(\bar{x}) > 0, \quad g(x) \in -C, \quad h(x) = 0_Z.$$

Moreover, we get for all $x \in S_d \times S_c$ with $x \notin S$ that $g(x) \notin -C$ or $h(x) \neq 0_Z$. Hence, the condition (6) is fulfilled.

(ii) Let $\bar{x}$ be no unique minimal solution of the discrete-continuous optimization problem (1). Then there is some $x \in S$ with $x \neq \bar{x}$ so that

$$f(x) - f(\bar{x}) \leq 0, \quad g(x) \in -C, \quad h(x) = 0_Z.$$

So, the condition (6) is not fulfilled.

(b) The proof of part (b) follows the lines of the proof of part (a). □

Next, we extend this result to problems without equality constraints and with a special ordering cone C.

Assumption 3.2. Let X_d and X_c be real linear spaces, and let Y be a real linear topological space. Let $C \subset Y$ be a closed convex cone with nonempty interior, and let there be a functional $\psi : Y \to \mathbb{R}$ with the properties (for arbitrary $y \in Y$)

$$\psi(y) = 0 \iff y \in \operatorname{bd}(-C)$$

and

$$\psi(y) < 0 \iff y \in \operatorname{int}(-C).$$

Let S_d be a nonempty discrete subset of X_d, and let S_c be a nonempty subset of X_c. Let $f : S_d \times S_c \to \mathbb{R}$ and $g : S_d \times S_c \to Y$ be given maps and let the constraint set

$$S := \{(x_d, x_c) \in S_d \times S_c \mid g(x_d, x_c) \in -C\}$$

be nonempty.

Remark 3.3. Notice that the functional ψ in Assumption 3.2 describes the shape of the closed convex cone $-C$, i.e. the level set of ψ at level 0 equals the boundary of $-C$. For an arbitrary $y \in Y$ we have with the closedness of C

$$\psi(y) \leq 0 \iff y \in \operatorname{int}(-C) \cup \operatorname{bd}(-C) = -C.$$

Such a functional is not unique (if ψ has the required properties, then $\alpha\psi$ also has these properties for every $\alpha > 0$).

For standard ordering cones we now present some associated functionals ψ.

Example 3.4.

(a) Let the real linear space $Y := \mathbb{R}^m$ (with $m \in \mathbb{N}$) be given.

(i) Consider a *polyhedral* closed convex cone with nonempty interior

$$C := \{y \in \mathbb{R}^m \mid a_i^T y \leq 0 \text{ for all } i \in \{1, \dots, k\}\}$$

for some number $k \in \mathbb{N}$ and nonzero vectors $a_1, \dots, a_k \in \mathbb{R}^m$. The functional $\psi : \mathbb{R}^m \to \mathbb{R}$ with

$$\psi(y) = \max_{i \in \{1,\dots,k\}} \{-a_i^T y\} \text{ for all } y \in \mathbb{R}^m$$

fulfills the required properties in Assumption 3.2. In the special case $C := \mathbb{R}^m_+$ it is evident that the functional ψ is given by

$$\psi(y_1, \dots, y_m) = \max\{y_1, \dots, y_m\} \text{ for all } (y_1, \dots, y_m) \in \mathbb{R}^m$$

(compare also [15, Example 1]). Figure 3.1 illustrates the cone $-\mathbb{R}^m_+$ and two level sets of the functional ψ for $m = 2$.

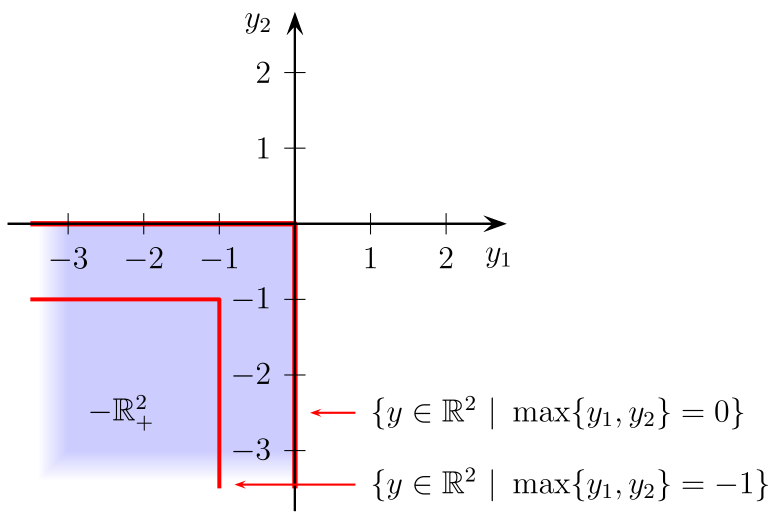

Figure 3.1: Illustration of level sets of the functional ψ in Example 3.4,(a),(i).

(ii) The functional $\psi : \mathbb{R}^m \to \mathbb{R}$ with

$$\psi(y) = \|(y_1, \dots, y_{m-1})\|_2 + y_m \text{ for all } y \in \mathbb{R}^m$$

(where $m \geq 2$ and $\|\cdot\|_2$ denotes the Euclidean norm in $\mathbb{R}^{m-1}$) is associated to the *Lorentz cone*

$$C := \{y \in \mathbb{R}^m \mid \|(y_1, \dots, y_{m-1})\|_2 \leq y_m \text{ for all } i \in \{1, \dots, k\}\}.$$

(b) Now consider the real linear space $Y := \mathcal{S}^m$ (with $m \in \mathbb{N}$) of all real symmetric (m, m) matrices. (i) If the well-known *Löwner cone*

$$C := \{M \in \mathcal{S}^m \mid y^T M y \geq 0 \text{ for all } y \in \mathbb{R}^m\} =: \mathcal{S}^m_+$$

is given, then the functional $\psi : \mathcal{S}^m \to \mathbb{R}$ with

$$\psi(M) = \max\{\lambda_1, \ldots, \lambda_m\} \text{ for all } M \in \mathcal{S}^m$$

is associated (here $\lambda_1, \ldots, \lambda_m$ denote the m eigenvalues of the matrix M).

(ii) For the *copositive cone*

$$C := \left\{M \in \mathcal{S}^m \;\middle|\; y^T M y \geq 0 \text{ for all } y \in \mathbb{R}^m_+\right\}$$

we obtain a functional $\psi : \mathcal{S}^m \to \mathbb{R}$ with

$$\psi(M) = \max_{y \in \mathbb{R}^m_+} y^T M y \text{ for all } M \in \mathcal{S}^m.$$

(c) Finally consider the infinite dimensional real linear space $Y := \mathcal{C}[a, b]$ of real-valued continuous functionals on $[a, b]$ with $-\infty < a < b < \infty$. For the natural ordering cone

$$C := \left\{y \in \mathcal{C}[a, b] \;\middle|\; y(t) \geq 0 \text{ for all } t \in [a, b]\right\}$$

we can work with a functional $\psi : \mathcal{C}[a, b] \to \mathbb{R}$ given by

$$\psi(y) = \max_{t \in [a,b]} y(t) \text{ for all } y \in \mathcal{C}[a, b]$$

(compare also [15, Example 3]).

Remark 3.5. There are close connections between the functional ψ and the so-called translation invariant functional (e.g. compare [13] and the references therein), and there are relationships to the so-called signed distance function due to Hiriart-Urruty (see [7] and compare [6, p. 439]). Although the formulas in Example 3.4 are simple to derive, the explicit formulation of ψ for a polyhedral closed convex cone with nonempty interior in Example 3.4,(a),(i) and the formula in Example 3.4,(c) can also be obtained as a special case by the translation invariant functional approach (see [13, Example 4.1.6]) and the signed-distance functional approach (see [15]).

Under Assumption 3.2 we now investigate the discrete-continuous optimization problem without equality constraint

$$\min_{(x_d, x_c) \in S} f(x_d, x_c). \tag{7}$$

Theorem 3.6. *Let Assumption* 3.2 *be satisfied, and let some* $\bar{x} \in S$ *be given.*
(a) *The vector* $\bar{x}$ *is a unique minimal solution of problem* (7) *if and only if*

$$\max\Big\{f(x) - f(\bar{x}),\ \psi\big(g(x)\big)\Big\} > 0 \text{ for all } x \in S_d \times S_c \text{ with } x \neq \bar{x}. \tag{8}$$

(b) *If the vector* $\bar{x}$ *is a minimal solution of problem* (7), *then*

$$\max\Big\{f(x) - f(\bar{x}),\ \psi\big(g(x)\big)\Big\} \geq 0 \text{ for all } x \in S_d \times S_c.$$

Proof.
(a) By Lemma 3.1,(a) $\bar{x}$ is a unique minimal solution of problem (7) if and only if

$$\big[-\mathbb{R}_+ \times (-C)\big] \cap \left\{ \left(\begin{array}{c} f(x) - f(\bar{x}) \\ g(x) \end{array} \right) \in \mathbb{R} \times Y \;\middle|\; x \in S_d \times S_c,\ x \neq \bar{x} \right\} = \emptyset. \quad (9)$$

Next, we prove for arbitrary $\alpha \in \mathbb{R}$ and $y \in Y$ the equivalence

$$(\alpha, y) \in -\mathbb{R}_+ \times (-C) \iff \max\{\alpha, \psi(y)\} \leq 0. \quad (10)$$

For the proof of this equivalence we first assume that $\max\{\alpha, \psi(y)\} \leq 0$. Then we get $\alpha \leq 0$ and $\psi(y) \leq 0$, i.e. $y \in -C$ (compare Remark 3.3), and we have $(\alpha, y) \in -\mathbb{R}_+ \times (-C)$. And conversely, assume that $(\alpha, y) \in -\mathbb{R}_+ \times (-C)$. Then we have $\alpha \leq 0$ and $\psi(y) \leq 0$ implying $\max\{\alpha, \psi(y)\} \leq 0$.

With the equivalence (10) together with the condition (9) we obtain that $\bar{x}$ is a unique minimal solution of problem (7) if and only if the condition (8) is fulfilled.

(b) Let $\bar{x}$ be a minimal solution of problem (7), and let $x \in S_d \times S_c$ be arbitrarily chosen. If $x \in S$, then $f(x) - f(\bar{x}) \geq 0$ and we get $\max\Big\{f(x) - f(\bar{x}),\ \psi\big(g(x)\big)\Big\} \geq 0$. If $x \notin S$, then we have $g(x) \notin -C$ being equivalent to $\psi\big(g(x)\big) > 0$ by Remark 3.3 and we conclude $\max\Big\{f(x) - f(\bar{x}),\ \psi\big(g(x)\big)\Big\} > 0$. This completes the proof. □

Remark 3.7.
(a) The optimality condition (8) in Theorem 3.6(a) is equivalent to the statement that $\bar{x}$ is a unique minimal solution of the optimization problem

$$\min_{x \in S_d \times S_c} \max\big\{f(x) - f(\bar{x}),\ \psi\big(g(x)\big)\big\}.$$

This problem has no inequality constraint, and in this sense it is "unconstrained".
(b) The proof of Theorem 3.6,(a) is based on a simple nonlinear separation of sets. In a more general framework such a kind of separation could also be achieved with [14, 13].

Next we specialize Assumption 3.2.

Assumption 3.8. Let X_d and X_c be real linear spaces, and let $Y := \mathbb{R}^m$ and $C := \mathbb{R}^m_+$ for some $m \in \mathbb{N}$ be given. Let S_d be a nonempty discrete subset of X_d, and let S_c be a nonempty subset of X_c. Let $f, g_1, \ldots, g_m : S_d \times S_c \to \mathbb{R}$ be given functionals and let the constraint set

$$S := \big\{(x_d, x_c) \in S_d \times S_c \mid g_k(x_d, x_c) \leq 0 \text{ for all } k \in \{1, \ldots, m\}\big\}$$

be nonempty.

Under this assumption the discrete-continuous optimization problem (7) reduces to the problem

$$\begin{aligned} &\min f(x_d, x_c) \;\text{ subject to the constraints} \\ &g_1(x_d, x_c) \le 0,\ \ldots,\ g_m(x_d, x_c) \le 0, \quad (x_d, x_c) \in S_d \times S_c. \end{aligned} \tag{11}$$

The following optimality conditions are simple to prove.

Lemma 3.9. *Let Assumption* 3.8 *be satisfied, and let some* $\bar{x} \in S$ *be given.*

(a) *The vector* $\bar{x}$ *is a unique minimal solution of problem* (11) *if and only if*

$$\max\left\{f(x) - f(\bar{x}),\ g_1(x),\ \ldots,\ g_m(x)\right\} > 0 \text{ for all } x \in S_d \times S_c \text{ with } x \ne \bar{x}. \tag{12}$$

(b) *If the vector* $\bar{x}$ *is a minimal solution of problem* (11)*, then*

$$\max\left\{f(x) - f(\bar{x}),\ g_1(x),\ \ldots,\ g_m(x)\right\} \ge 0 \text{ for all } x \in S_d \times S_c.$$

Remark 3.10.

(a) Although Lemma 3.9 is simple to prove, part (a) is also a direct consequence of Theorem 3.6,(a) and Example 3.4,(a),(i), if we notice that for arbitrary $x \in S_d \times S_c$ with $x \ne \bar{x}$

$$\begin{aligned} &\max\Big\{f(x) - f(\bar{x}),\ \max\left\{g_1(x),\ \ldots,\ g_m(x)\right\}\Big\} > 0 \\ \Longleftrightarrow\ &\max\Big\{f(x) - f(\bar{x}),\ g_1(x),\ \ldots,\ g_m(x)\Big\} > 0. \end{aligned}$$

(b) In analogy to Remark 3.7,(a) the optimality condition (12) in Lemma 3.9,(a) is equivalent to the statement that $\bar{x}$ is a unique minimal solution of the optimization problem

$$\min_{x \in S_d \times S_c} \max\left\{f(x) - f(\bar{x}),\ g_1(x),\ \ldots,\ g_m(x)\right\}. \tag{13}$$

(c) Notice that the optimality condition for a minimal solution in part (b) of Lemma 3.9 is only a necessary condition. In general, the sufficiency of this condition does not hold.

For simplicity the objective function of problem (13) is denoted by $\varphi : S_d \times S_c \to \mathbb{R}$ with

$$\varphi(x) = \max\left\{f(x) - f(\bar{x}),\ g_1(x),\ \ldots, g_m(x)\right\} \text{ for all } x \in S_d \times S_c.$$

We illustrate Remark 3.10,(b) with a very simple example.

Example 3.11. We investigate the discrete-continuous optimization problem

$$\begin{aligned} &\min\ \left(x_d - \tfrac{1}{2}\right)^2 + x_c^2 \\ &\text{subject to the constraint} \quad -x_d + \tfrac{3}{4} \le 0, \quad x_d \in \mathbb{Z},\ x_c \in \mathbb{R}. \end{aligned} \tag{14}$$

It is evident that $\bar{x} := (1, 0)^T$ is the unique minimal solution of problem (14). Then the problem (13) can be written as

$$\min_{(x_d, x_c) \in \mathbb{Z} \times \mathbb{R}} \underbrace{\max\left\{\left(x_d - \tfrac{1}{2}\right)^2 + x_c^2 - \tfrac{1}{4},\ -x_d + \tfrac{3}{4}\right\}}_{=\varphi(x_d, x_c)}. \tag{15}$$

For an arbitrary $x_c \in \mathbb{R}$ we then obtain

$$\begin{aligned}
&\varphi(-2,x_c) = \max\left\{6+x_c^2,\ \tfrac{11}{4}\right\} = 6+x_c^2 > 0\\
&\varphi(-1,x_c) = \max\left\{2+x_c^2,\ \tfrac{7}{4}\right\} = 2+x_c^2 > 0\\
&\varphi(0,x_c) = \max\left\{x_c^2,\ \tfrac{3}{4}\right\} > 0\\
&\varphi(1,x_c) = \max\left\{x_c^2,\ -\tfrac{1}{4}\right\} = x_c^2 > 0 \text{ (for } x_c \neq 0)\\
&\varphi(2,x_c) = \max\left\{2+x_c^2,\ -\tfrac{5}{4}\right\} = 2+x_c^2 > 0\\
&\varphi(3,x_c) = \max\left\{6+x_c^2,\ -\tfrac{9}{4}\right\} = 6+x_c^2 > 0
\end{aligned}$$

and so on. Consequently, $\bar{x}$ is a unique minimal solution of the unconstrained problem (15).

Next, we apply the theory of subdifferentials, which has to be adapted to the case of unique minimal solutions and discrete-continuous variables.

Definition 3.12. Let X_d and X_c be real linear spaces. Let S_d be a nonempty discrete subset of X_d, and let S_c be a nonempty subset of X_c. Let $\xi : X_d \times X_c \to \mathbb{R}$ be a given functional. For an arbitrary $\bar{x} \in S_d \times S_c$ the set

$$\partial_{>}\xi(\bar{x}) := \left\{\ell \in (X_d \times X_c)' \;\middle|\; \xi(x) > \xi(\bar{x}) + \ell(x-\bar{x}) \text{ for all } x \in S_d \times S_c \text{ with } x \neq \bar{x}\right\}$$

is called the *strict subdifferential* of ξ at $\bar{x}$.

Remark 3.13.

(a) In addition to the assumptions of Definition 3.12 let $(X_d, \|\cdot\|_{X_d})$ and $(X_c, \|\cdot\|_{X_c})$ be real normed spaces and let the functional ξ be Gâteaux differentiable at $\bar{x}$ and strictly convex at $\bar{x}$, i.e. for an arbitrary $x \in S_d \times S_c$, $x \neq \bar{x}$,

$$\xi(\lambda x + (1-\lambda)\bar{x}) < \lambda\xi(x) + (1-\lambda)\xi(\bar{x}) \text{ for all } \lambda \in (0,1).$$

Then we get with the Gâteaux derivative $\xi'(\bar{x})(\cdot)$ of ξ at $\bar{x}$

$$\xi(x) > \xi(\bar{x}) + \xi'(\bar{x})(x-\bar{x}) \text{ for all } x \in S_d \times S_c \text{ with } x \neq \bar{x}$$

(compare [3, Satz 4] together with [8, Thm. 3.16]), i.e. we have $\xi'(\bar{x}) \in \partial_{>}\xi(\bar{x})$, and thus $\partial_{>}\xi(\bar{x}) \neq \emptyset$.

(b) In addition to the assumptions of Definition 3.12 let $\bar{x} = (\bar{x}_d, \bar{x}_c) \in S_d \times S_c$ with $\bar{x}_c \in \operatorname{cor}(S_c)$ and consider an affine linear functional ξ, i.e. for some linear functional $\bar{\ell} \in (X_d \times X_c)'$ and some $\alpha \in \mathbb{R}$ we have

$$\xi(x) = \bar{\ell}(x) + \alpha \text{ for all } x \in S_d \times S_c.$$

Then we obtain $\xi(x) - \xi(\bar{x}) = \bar{\ell}(x-\bar{x})$ for all $x \in S_d \times S_c$.

If we assume that there is some $\ell \in \partial_{>}\xi(\bar{x})$, we conclude

$$\left(\bar{\ell}-\ell\right)(x-\bar{x}) > 0 \text{ for all } x \in S_d \times S_c \text{ with } x \neq \bar{x}.$$

And since $\bar{x}_c \in \operatorname{cor}(S_c)$, there is some vector $h_c \in X_c \backslash \{0_{X_c}\}$ with $\bar{x}_c \pm h_c \in S_c$ so that we get a contradiction with

$$0 < \left(\bar{\ell}-\ell\right)\left((\bar{x}_d, \bar{x}_c + h_c) - (\bar{x}_d, \bar{x}_c)\right) = \left(\bar{\ell}-\ell\right)(0_{X_d}, h_c)$$

and

$$0 < \left(\bar{\ell}-\ell\right)\left((\bar{x}_d, \bar{x}_c - h_c) - (\bar{x}_d, \bar{x}_c)\right) = -\left(\bar{\ell}-\ell\right)(0_{X_d}, h_c).$$

Consequently, the strict subdifferential of an affine linear functional is empty at every $\bar{x} = (\bar{x}_d, \bar{x}_c) \in S_d \times S_c$ with $\bar{x}_c \in \operatorname{cor}(S_c)$.

We now present a sufficient optimality condition without any explicit convexity assumption.

Theorem 3.14. *Let Assumption* 3.8 *be satisfied, and let $\bar{x} \in S$ be a feasible point. Let*

$$I(\bar{x}) := \big\{k \in \{1, \dots, m\} \;\big|\; g_k(\bar{x}) = 0\big\} \tag{16}$$

denote the index set of "active" inequality constraints, and let

$$J(\bar{x}) := \big\{k \in \{1, \dots, m\} \backslash I(\bar{x}) \;\big|\; g_k(x) \le g_k(\bar{x}) \text{ for all } x \in S\big\} \tag{17}$$

denote the index set of "nearly active" inequality constraints. Let there exist nonnegative real numbers α, β_k $(k \in I(\bar{x}))$, γ_k $(k \in J(\bar{x}))$, where at least one of them is nonzero, and linear functionals $\ell_0 \in \partial_> f(\bar{x})$, $\ell_k \in \partial_> g_k(\bar{x})$ $(k \in I(\bar{x}))$, $\ell_k \in \partial_> g_k(\bar{x})$ $(k \in J(\bar{x}))$ so that

$$\alpha \ell_0 + \sum_{k \in I(\bar{x})} \beta_k \ell_k + \sum_{k \in J(\hat{x})} \gamma_k \ell_k = 0_{(X_d \times X_c)'}. \tag{18}$$

Then the vector $\bar{x}$ is a unique minimal solution of the discrete-continuous optimization problem (11).

Proof. Let the given assumptions be satisfied. With the equality (18) we then obtain for arbitrary $x \in S_d \times S_c$ with $x \neq \bar{x}$

$$\begin{aligned}
&\underbrace{\max\big\{f(x) - f(\bar{x}),\ g_1(x),\ \dots,\ g_m(x),\ (g_k(x) - g_k(\bar{x}))_{k \in J(\bar{x})}\big\}}_{=:\hat{\varphi}(x)} \\
&= \underbrace{\frac{1}{\alpha + \sum_{k \in I(\bar{x})} \beta_k + \sum_{k \in J(\bar{x})} \gamma_k}}_{=:\rho > 0} \Big(\alpha \hat{\varphi}(x) + \sum_{k \in I(\bar{x})} \beta_k \hat{\varphi}(x) + \sum_{k \in J(\bar{x})} \gamma_k \hat{\varphi}(x)\Big) \\
&\ge \rho \Big(\alpha\big(f(x) - f(\bar{x})\big) + \sum_{k \in I(\bar{x})} \beta_k g_k(x) + \sum_{k \in J(\bar{x})} \gamma_k \big(g_k(x) - g_k(\bar{x})\big)\Big) \\
&> \rho \Big(\alpha \ell_0(x - \bar{x}) + \sum_{k \in I(\bar{x})} \beta_k \Big[\underbrace{g_k(\bar{x})}_{=0} + \ell_k(x - \bar{x})\Big] + \sum_{k \in J(\bar{x})} \gamma_k \ell_k(x - \hat{x})\Big) \\
&= \rho \Big(\alpha \ell_0 + \sum_{k \in I(\bar{x})} \beta_k \ell_k + \sum_{k \in J(\bar{x})} \gamma_k \ell_k\Big)(x - \bar{x}) \\
&= 0.
\end{aligned}$$

With Lemma 3.9(a) the vector $\bar{x}$ is a unique minimal solution of the problem

$$\begin{gathered}
\min f(x) \ \text{ subject to the constraints} \\
g_1(x) \le 0,\ \dots\ , g_m(x) \le 0, \\
g_k(x) - g_k(\bar{x}) \le 0 \text{ for all } k \in J(\bar{x}),\quad x \in S_d \times S_c
\end{gathered} \tag{19}$$

Next, we show that the inequality constraints in problem (19) defined by the index set $J(\bar{x})$ are redundant, i.e.

$$S = \{x \in S \mid g_k(x) - g_k(\bar{x}) \leq 0 \text{ for all } k \in J(\bar{x})\}. \tag{20}$$

It is obvious that the right hand side in (20) is a subset of S. For the proof of the converse inclusion choose an arbitrary $x \in S$, and we obtain with the definition in (17)

$$g_k(x) - g_k(\bar{x}) \leq 0 \text{ for all } k \in J(\bar{x}). \tag{21}$$

Hence, the equation (20) is shown and the inequalities (21) can be dropped in problem (19). Consequently, the vector $\bar{x}$ is a unique minimal solution of the discrete-continuous optimization problem (11). □

Remark 3.15.

(a) If the discrete-continuous optimization problem (11) is unconstrained, i.e. there are no inequality constraints, then the equality (18) reduces to $\ell_0 = 0_{(X_d \times X_c)'}$ and we obtain $0_{(X_d \times X_c)'} \in \partial_> f(\bar{x})$. Obviously, this is a sufficient condition for the unique minimality of $\bar{x}$.

(b) Notice that the equality (18) with nonnegative real numbers α, β_k ($k \in I(\bar{x})$), γ_k ($k \in J(\bar{x})$), where at least one of them is nonzero, and linear functionals $\ell_0 \in \partial_> f(\bar{x})$, $\ell_k \in \partial_> g_k(\bar{x})$ ($k \in I(\bar{x})$), $\ell_k \in \partial_> g_k(\bar{x})$ ($k \in J(\bar{x})$) implies

$$0_{(X_d \times X_c)'} \in \operatorname{conv}\Bigg(\partial_> f(\bar{x}) \cup \bigcup_{k \in I(\bar{x})} \partial_> g_k(\bar{x}) \cup \bigcup_{k \in J(\bar{x})} \partial_> g_k(\bar{x})\Bigg).$$

The first part of this condition is well-known from continuous optimization (e.g. compare [11, Prop. 2.54]), and the part with the index set $J(\bar{x})$ is needed for discrete variables.

(c) The index set $J(\bar{x})$ is needed in Theorem 3.14 because the notion of active inequality constraints is generally too strict in discrete optimization. We have to incorporate inequality constraints, which are "nearly active". The involvement of the index set $J(\bar{x})$ in the optimality condition (18) is one possibilty to handle discrete-continuous problems. The following example illustrates this point.

(d) The optimality conditions in Theorem 3.14 extend the standard Fritz John conditions [10] to discrete-continuous optimization problems.

Example 3.16. We modify the discrete-continuous optimization problem (14) in Example 3.11 and we investigate the problem

$$\begin{gathered} \min \ \left(x_d - \tfrac{1}{2}\right)^2 + x_c^2 \ \text{ subject to the constraint} \\ (x_d - 5)^2 - 17 \leq 0, \quad x_d \in \mathbb{Z},\ x_c \in \mathbb{R}. \end{gathered} \tag{22}$$

In this way the constraint function has a nonempty strict subdifferential. With the objective function $f : \mathbb{Z} \times \mathbb{R} \to \mathbb{R}$ given by

$$f(x_d, x_c) = \left(x_d - \tfrac{1}{2}\right)^2 + x_c^2 \text{ for all } x_d \in \mathbb{Z} \text{ and } x_c \in \mathbb{R}$$

and the constraint function $g_1 : \mathbb{Z} \times \mathbb{R} \to \mathbb{R}$ defined by

$$g_1(x_d, x_c) = (x_d - 5)^2 - 17 \text{ for all } x_d \in \mathbb{Z} \text{ and } x_c \in \mathbb{R}$$

we obtain for arbitrary $x_d \in \mathbb{Z}$ and $x_c \in \mathbb{R}$

$$\partial_> f(x_d, x_c) = \{\nabla f(x_d, x_c)\} = \left\{ \begin{pmatrix} 2(x_d - \frac{1}{2}) \\ 2x_c \end{pmatrix} \right\}$$

and
$$\partial_> g_1(x_d, x_c) = \{\nabla g_1(x_d, x_c)\} = \left\{ \begin{pmatrix} 2(x_d - 5) \\ 0 \end{pmatrix} \right\}.$$

Now we set $\bar{x} := (1, 0)^T \in S$ and we get in consequence $\ell_0 := \nabla f(\bar{x}) = (1, 0)^T$ and $\ell_1 := \nabla g_1(\bar{x}) = (-8, 0)^T$. The inequality constraint in problem (22) is not "active", i.e. $I(\bar{x}) = \emptyset$, but it is "nearly active", i.e. $J(\bar{x}) = \{1\}$, because the feasible set is given by $S := \{1, 2, \ldots, 9\} \times \mathbb{R}$ and we have

$$g_1(x) \leq -1 = g_1(\bar{x}) \text{ for all } x \in S.$$

With $\alpha := \frac{8}{9}$ and $\beta_1 := \frac{1}{9}$ we get

$$\alpha\ell_0 + \beta_1\ell_1 = \frac{8}{9}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{9}\begin{pmatrix} -8 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

i.e., the equality (18) is fulfilled. With Theorem 3.14 we then conclude that $\bar{x}$ is a unique minimal solution of the discrete-continuous optimization problem (22).

For the proof of a necessary optimality condition for a unique minimal solution $\bar{x}$ of the discrete-continuous optimization problem (11) we define an index set $K(\bar{x}) \subset \{1, \ldots, m\}$ with the smallest magnitude so that $\bar{x}$ is also a unique minimal solution of the problem

$$\begin{gathered} \min f(x) \quad \text{subject to the constraints} \\ g_k(x) \leq 0 \text{ for all } k \in K(\bar{x}), \quad x \in S_d \times S_c. \end{gathered} \tag{23}$$

Hence, for $k \in K(\bar{x})$ the k-th constraint is not redundant. Notice that $K(\bar{x})$ is generally not uniquely defined. The set

$$\hat{I}(\bar{x}) := \{k \in K(\bar{x}) \mid g_k(\bar{x}) = 0\}$$

denotes the set of "active" inequality constraints among the constraints with index in $K(\bar{x})$. Moreover, we use the index set $\hat{J}(\bar{x}) := K(\bar{x}) \backslash \hat{I}(\bar{x})$, which identifies non-active constraints.

In the following theorem we use the abbreviation $\mathcal{O}$ for the "null" functional $\mathcal{O} : S_d \times S_c \to \mathbb{R}$ with $\mathcal{O}(x) = 0$ for all $x \in S_d \times S_c$.

Theorem 3.17. *Let Assumption* 3.8 *be satisfied. If the vector $\bar{x}$ is a unique minimal solution of the discrete-continuous optimization problem* (11) *and $K(\bar{x})$ is an index set as defined above, then*

$$0_{(X_d \times X_c)'} \in \partial_> \max\left\{f(\cdot) - f(\bar{x}),\ (g_k)_{k \in \hat{I}(\bar{x})},\ (g_k(\cdot) - g_k(\bar{x}))_{k \in \hat{J}(\bar{x})}\right\}\Big|_{x=\bar{x}}. \tag{24}$$

If, in addition, the "intersection" rule for normal cones

$$\begin{aligned} N\Big(&\operatorname{epi}\big(\max\big\{f(\cdot)-f(\bar{x}),\ (g_k)_{k\in\hat{I}(\bar{x})},\ (g_k(\cdot)-g_k(\bar{x}))_{k\in\hat{J}(\bar{x})}\big\}\big),(\bar{x},0)\Big) \\ &= N\big(\operatorname{epi}(f(\cdot)-f(\bar{x})),(\bar{x},0)\big)+\sum_{k\in\hat{I}(\bar{x})} N\big(\operatorname{epi}(g_k),(\bar{x},0)\big) \\ &\quad+\sum_{k\in\hat{J}(\bar{x})} N\big(\operatorname{epi}(g_k(\cdot)-g_k(\bar{x})),(\bar{x},0)\big) \end{aligned} \tag{25}$$

is fulfilled and the subdifferentials $\partial f(\bar{x})$ and $\partial g_k(\bar{x})$ ($k\in K(\bar{x})$) are nonempty, then there are positive real numbers α, β_k ($k\in K(\bar{x})$), and linear functionals $\ell_0\in\partial f(\bar{x})\cup\partial\mathcal{O}(\bar{x})$, $\ell_k\in\partial g_k(\bar{x})\cup\partial\mathcal{O}(\bar{x})$ ($k\in K(\bar{x})$), where at least one of these linear functionals only belongs to $\partial f(\bar{x})$ and $\partial g_k(\bar{x})$ ($k\in K(\bar{x})$), respectively, with the property

$$\alpha\ell_0+\sum_{k\in K(\bar{x})}\beta_k\ell_k=0_{(X_d\times X_c)'}. \tag{26}$$

Proof. Let $\bar{x}$ be a unique minimal solution of the discrete-continuous optimization problem (11). By the definition of $K(\bar{x})$ the vector $\bar{x}$ is a unique minimal solution of the problem (23). We modify problem (23) and we investigate the optimization problem

$$\begin{gathered} \min f(x) \\ \text{subject to the constraints } \ g_k(x)\le 0 \text{ for all } k\in\hat{I}(\bar{x}), \\ g_k(x)-g_k(\bar{x})\le 0 \text{ for all } k\in\hat{J}(\bar{x}),\quad x\in S_d\times S_c. \end{gathered} \tag{27}$$

Since $\bar{x}$ is a feasible point of problem (27) and the set of feasible points of problem (27) is contained in the set of feasible points of problem (23), $\bar{x}$ is a unique minimal solution of problem (27) as well. With Lemma 3.9,(a) we then obtain for the functional $\hat{\varphi}:S_d\times S_c\to\mathbb{R}$ with

$$\hat{\varphi}(x):=\max\big\{f(x)-f(\bar{x}),\ (g_k(x))_{k\in\hat{I}(\bar{x})},\ (g_k(x)-g_k(\bar{x}))_{k\in\hat{J}(\bar{x})}\big\} \text{ for all } x\in S_d\times S_c$$

the inequality

$$\hat{\varphi}(x)>0=\hat{\varphi}(\bar{x}) \text{ for all } x\in S_d\times S_c \text{ with } x\ne\bar{x}. \tag{28}$$

Then the optimality condition $0_{(X_d\times X_c)'}\in\partial_>\hat{\varphi}(\bar{x})$ is fulfilled, i.e. the condition (24) is shown.

Now we prove the second part of the assertion. Since the "standard" subdifferential $\partial\hat{\varphi}(\bar{x})$ is a superset of the strict subdifferential $\partial_>\hat{\varphi}(\bar{x})$, it follows from (24) $0_{(X_d\times X_c)'}\in\partial\hat{\varphi}(\bar{x})$. For an arbitrary $(x,\lambda)\in\operatorname{epi}(\hat{\varphi})$, i.e. for arbitrary $x\in S_d\times S_c$ and $\lambda\ge\hat{\varphi}(x)$, we conclude with (28)

$$\big(0_{(X_d\times X_c)'},-1\big)\big((x,\lambda)-(\bar{x},\hat{\varphi}(\bar{x}))\big)=-\lambda+\hat{\varphi}(\bar{x})\le-\hat{\varphi}(x)+\hat{\varphi}(\bar{x})\le 0$$

(e.g. compare the proof of [11, Prop. 2.31]). This means

$$\big(0_{(X_d\times X_c)'},\ -1\big)\in N\big(\operatorname{epi}(\hat{\varphi}),(\bar{x},\hat{\varphi}(\bar{x}))\big)=N\big(\operatorname{epi}(\hat{\varphi}),(\bar{x},0)\big)$$

and with the "intersection" rule (25) for normal cones we obtain

$$\begin{aligned}\big(0_{(X_d\times X_c)'},\ -1\big) \ \ \in \ \ & N\big(\mathrm{epi}(f(\cdot)-f(\bar{x})),(\bar{x},0)\big) + \sum_{k\in\hat{I}(\bar{x})} N\big(\mathrm{epi}(g_k),(\bar{x},0)\big)\\ & + \sum_{k\in\hat{J}(\bar{x})} N\big(\mathrm{epi}(g_k(\cdot)-g_k(\bar{x})),(\bar{x},0)\big).\end{aligned}$$

Hence, there are elements

$$(\bar{\ell}_0,-\bar{\alpha}) \in N\big(\mathrm{epi}(f(\cdot)-f(\bar{x})),(\bar{x},0)\big), \quad (\bar{\ell}_k,-\bar{\beta}_k) \in N\big(\mathrm{epi}(g_k),(\bar{x},0)\big)\ (k\in\hat{I}(\bar{x}))$$

and $(\bar{\ell}_k,-\bar{\beta}_k) \in N\big(\mathrm{epi}(g_k(\cdot)-g_k(\bar{x})),(\bar{x},0)\big)$ $(k\in\hat{J}(\bar{x}))$ with

$$\big(0_{(X_d\times X_c)'},\ -1\big) = (\bar{\ell}_0,-\bar{\alpha}) + \sum_{k\in\hat{I}(\bar{x})}(\bar{\ell}_k,-\bar{\beta}_k) + \sum_{k\in\hat{J}(\bar{x})}(\bar{\ell}_k,-\bar{\beta}_k). \tag{29}$$

The condition $(\bar{\ell}_0,-\bar{\alpha}) \in N\big(\mathrm{epi}(f(\cdot)-f(\bar{x})),(\bar{x},0)\big)$ means that

$$\bar{\ell}_0(x-\bar{x})-\bar{\alpha}\lambda \le 0 \text{ for all } (x,\lambda)\in \mathrm{epi}(f(\cdot)-f(\bar{x})). \tag{30}$$

For $x=\bar{x}$ and $\lambda=1$ this inequality implies $\bar{\alpha}\ge 0$. If $\bar{\alpha}=0$, the inequality (30) yields

$$\bar{\ell}_0(x-\bar{x}) \le 0 \text{ for all } x\in S_d\times S_c,$$

i.e. $\bar{\ell}_0\in\partial\mathcal{O}(\bar{x})$. If $\bar{\alpha}>0$, it follows from the inequality (30)

$$\bar{\alpha}f(x) \ge \bar{\alpha}f(\bar{x}) + \bar{\ell}_0(x-\bar{x}) \text{ for all } x\in S_d\times S_c,$$

i.e. $\bar{\ell}_0\in\bar{\alpha}\partial f(\bar{x})$. Consequently, we have

$$\bar{\ell}_0 = \alpha\ell_0 \text{ for some } \alpha>0 \text{ and some } \ell_0\in\partial f(\bar{x})\cup\partial\mathcal{O}(\bar{x}).$$

In a similar way we get for all $k\in\hat{I}(\bar{x})\cup\hat{J}(\bar{x})$ that $\bar{\beta}_k\ge 0$ and that

$$\bar{\ell}_k = \beta_k\ell_k \text{ for some } \beta_k>0 \text{ and some } \ell_k\in\partial g_k(\bar{x})\cup\partial\mathcal{O}(\bar{x}).$$

Then it follows from the equality (29)

$$0_{(X_d\times X_c)'} = \alpha\ell_0 + \sum_{k\in\hat{I}(\bar{x})}\beta_k\ell_k + \sum_{k\in\hat{J}(\bar{x})}\beta_k\ell_k \tag{31}$$

and

$$-1 = -\bar{\alpha} - \sum_{k\in\hat{I}(\bar{x})}\bar{\beta}_k - \sum_{k\in\hat{J}(\bar{x})}\bar{\beta}_k. \tag{32}$$

The equality (32) means that at least one of the parameters is positive or, in other words, at least one of the subgradients in the equality (31) only belongs to $\partial f(\bar{x})$ and $\partial g_k(\bar{x})$ $(k\in K(\bar{x}))$, respectively. If we notice that $\hat{J}(\bar{x}) = K(\bar{x})\backslash\hat{I}(\bar{x})$ the equality (31) implies the equality (26), which has to be shown. □

Remark 3.18.

(a) The validity of the "intersection" rule (25) in the prevoius theorem is a strong assumption. In Banach spaces Burachik and Jeyakumar [2, Thm. 3.1] present assumptions for convex epigraphs, under which the intersection rule holds. For non-convex epigraphs, which may occur in discrete optimization, the intersection rule

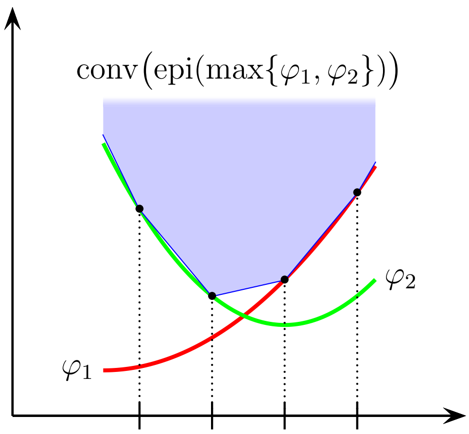

Figure 3.2: Illustration of the convex hull of the epigraph of the max functional with discrete preimages (compare Remark 3.18,(a)).

may not be true. As a simple example consider in Figure 3.2 the two functionals φ_1 and φ_2 defined on a discrete set. In this case we have

$$\operatorname{conv}\big(\operatorname{epi}(\max\{\varphi_1, \varphi_2\})\big) \neq \operatorname{conv}\big(\operatorname{epi}(\varphi_1)\big) \cap \operatorname{conv}\big(\operatorname{epi}(\varphi_2)\big)$$

so that the convex hull operation is not helpful even if φ_1 and φ_2 are convex.

(b) In analogy to Remark 3.15,(c) the necessary optimality conditions in the previous theorem extend the standard Fritz John conditions [10] to discrete-continuous optimization problems.

Conclusion

This new general nonlinear separation approach for optimality conditions is specialized to problems with a finite number of inequality constraints with values in $\mathbb{R}$. Based on Example 3.4 this theory could be extended to inequalities in the real linear space $\mathcal{S}^m$ and infinite dimensional linear spaces. Moreover, instead of a subdifferential approach one could also work with different nonsmooth techniques. But these are topics for future investigations.

Acknowledgement

The second author thanks Petra Weidner for providing helpful bibliographical references for this paper.

References

[1] M. Baes, T. Oertel, R. Weismantel: *Duality for mixed-integer convex minimization*, Math. Program., Ser. A 158 (2016) 547–564.

[2] R. S. Burachik, V. Jeyakumar: *A simple closure condition for the normal cone intersection formula*, Proc. Amer. Math. Soc. 133 (2004) 1741–1748.

[3] L. Collatz, W. Wetterling: *Optimierungsaufgaben*, Springer, Berlin (1971).

[4] J.-P. Doignon: *Convexity in cristallographical lattices*, J. Geom. 3 (1973) 71–85.

[5] C. Gerth, P. Weidner: *Nonconvex separation theorems and some applications in vector optimization*, J. Optimization Theory Appl. 67 (1990) 297–320.

[6] T. X. D. Ha: *The Fermat rule and Lagrange multiplier rule for various efficient solutions of set-valued optimization problems expressed in terms of coderivatives*, in: *Recent Developments in Vector Optimization*, Chapter 12, Q. H. Ansari, J.-C. Yao (eds.), Springer, Berlin (2012).

[7] J. B. Hiriart-Urruty: *New concepts in nondifferentiable programming*, Bull. Soc. Math. France 60 (1979) 57–85.

[8] J. Jahn: *Introduction to the Theory of Nonlinear Optimization*, Springer, Cham (2020).

[9] J. Jahn, M. Knossalla: *Lagrange theory of discrete-continuous nonlinear optimization*, J. Nonlinear Var. Analysis 2 (2018) 317–342.

[10] F. John: *Extremum problems with inequalities as subsidiary conditions*, in: *Studies and Essays*, K. O. Friedrichs, O. E. Neugebauer, J. J. Stoker (eds.), Courant Anniversary Volume, Interscience, New York (1948) 187–204.

[11] B. S. Mordukhovich, N. M. Nam: *An Easy Path to Convex Analysis and Applications*, Morgan & Claypool, Williston (2014).

[12] C. Planiden, X. Wang: *Most convex functions have unique minimizers*, J. Convex Analysis 23 (2016) 877–892.

[13] C. Tammer. P. Weidner: *Scalarization and Separation by Translation Invariant Functions - with Applications in Optimization, Nonlinear Functional Analysis, and Mathematical Economics*, Springer, Cham (2020).

[14] P. Weidner: *Lower semicontinuous functionals with uniform sublevel sets*, Optimization 66 (2017) 491–505.

[15] A. Zaffaroni: *Degrees of efficiency and degrees of minimality*, SIAM J. Control Optimization 42 (2003) 1071–1086.