Journal of Convex Analysis
Volume 11 (2004), No. 1, 95–110

On the Necessity of some Constraint Qualification Conditions in Convex Programming

Dan Tiba
Weierstrass Institut, Mohrenstr. 39, D-10117 Berlin, Germany
tiba@wias-berlin.de
permanent address: Institute of Mathematics, Romanian Academy,
P.O. Box 1-764, RO–70700, Bucharest, Romania
dtiba@imar.ro

Constantin Zălinescu
University "Al. I. Cuza" Iaşi, Faculty of Mathematics,
Bd. Copou Nr. 11, 6600 Iaşi, Romania
zalinesc@uaic.ro

Received February 18, 2002
Revised manuscript received December 11, 2002

In this paper we realize a study of various constraint qualification conditions for the existence of Lagrange multipliers for convex minimization problems in general normed vector spaces; it is based on a new formula for the normal cone to the constraint set, on local metric regularity and a metric regularity property on bounded subsets. As a by-product we obtain a characterization of the metric regularity of a finite family of closed convex sets.

Keywords: Convex function, constraint qualification, Lagrange multiplier, metric regularity, normal cone

2000 Mathematics Subject Classification: 49K27, 90C25

1. Introduction

Consider the classical convex programming problem

$$\text{minimize} \quad g(x) \quad \text{s.t.} \quad h_i(x) \le 0,\ i \in I := \{1, \ldots, m\}, \tag{P}$$

where g and h_i $(i \in I)$ are convex functions defined on the normed vector space X. We are interested by the weakest hypotheses that ensure the characterization of a solution $\overline{x}$ of (P) by Karush-Kuhn-Tucker conditions, *i.e.*, the existence of $\lambda_1, \ldots, \lambda_m \ge 0$, called Lagrange multipliers, such that $\overline{x}$ is a solution of the (unconstrained) minimization problem

$$\text{minimize} \quad g(x) + \sum\nolimits_{i=1}^{m} \lambda_i h_i(x) \quad \text{s.t.} \quad x \in X \tag{UP}$$

and $\lambda_i h_i(\overline{x}) = 0$ for every $i \in I$.

There are several known assumptions of this type in the literature, called constraint qualification (CQ) conditions. The mostly used seems to be Slater's CQ:

$$\exists\, \widetilde{x},\ \forall\, i \in I \ :\ h_i(\widetilde{x}) < 0. \tag{SCQ}$$

Denoting by C the set $\{x \in X \mid h_i(x) \le 0\ \forall\, i \in I\}$ of admissible solutions of (P) and by $I(\overline{x})$ the set $\{i \in I \mid h_i(\overline{x}) = 0\}$ of active constraints at $\overline{x} \in C$, and assuming the functions

ISSN 0944-6532 / \$ 2.50

h_i to be finite-valued and differentiable, other conditions are:

$$\{\nabla h_i(\overline{x}) \mid i \in I(\overline{x})\} \quad \text{is linearly independent,} \tag{LICQ}$$

the Mangasarian-Fromovitz' CQ

$$\exists\, \widetilde{u},\ \forall\, i \in I(\overline{x}) \ :\ \nabla h_i(\overline{x})(\widetilde{u}) < 0, \tag{MFCQ}$$

or Abadie's CQ

$$\overline{\operatorname{cone}}(C - \overline{x}) = \{u \mid \nabla h_i(\overline{x})(u) \leq 0\ \forall\, i \in I(\overline{x})\}. \tag{ACQ}$$

As noted by Li [19], (ACQ) is equivalent to the condition

$$N(C, \overline{x}) = \left\{ \sum\nolimits_{i \in I(\overline{x})} \lambda_i \nabla h_i(\overline{x}) \,\middle|\, \lambda_i \geq 0\ \forall\, i \in I(\overline{x}) \right\}, \tag{ACQ$'$}$$

where $N(C, \overline{x})$ is the normal cone of C at $\overline{x}$. Abadie's CQ is, consequently, a particular case of the "basic" constraint qualification introduced by Hiriart-Urruty and Lemaréchal [15] in the case of nondifferentiable convex minimization problems:

$$N(C, \overline{x}) = \left\{ \sum\nolimits_{i \in I(\overline{x})} \lambda_i x_i^* \,\middle|\, \lambda_i \geq 0,\ x_i^* \in \partial h_i(\overline{x})\ \forall\, i \in I(\overline{x}) \right\}. \tag{BCQ}$$

It is known that in finite dimensional spaces (see [15, 19, 8]) (LICQ) $\Rightarrow$ (SCQ) $\Leftrightarrow$ (MFCQ) $\Rightarrow$ (ACQ) $\Leftrightarrow$ (BCQ) when the functions h_i are convex and differentiable and (SCQ) $\Rightarrow$ (BCQ) when the functions h_i are finite-valued continuous and convex. As proved by Hiriart-Urruty and Lemaréchal [15] in finite dimensional spaces and for finite-valued convex functions, (BCQ) is also necessary for the existence of Lagrange multipliers in the sense that (BCQ) holds for $h_1, \ldots, h_m$ if for every continuous convex objective function g and any solution $\overline{x}$ of (P) there exist Lagrange multipliers $\lambda_1, \ldots, \lambda_m \geq 0$. In this sense, (BCQ) is the weakest possible CQ, but it is difficult to be checked due to its implicit character.

The aim of this article is to introduce another CQ, which is formulated directly in terms of the data:

$$\forall\, B \text{ bounded},\ \exists\, \gamma_B > 0,\ \forall\, x \in B \setminus C \ :\ \max\{h_i(x) \mid i \in I\} \geq \gamma_B \cdot d_C(x). \tag{MRB}$$

We show that (MRB) is strictly weaker than Slater's CQ, and is also necessary for the existence of Lagrange multipliers in the sense mentioned above, at least in finite dimensional spaces. This condition, which is a metric regularity condition on bounded sets, was used in another context by Robinson [24] and by Lemaire [18]. Meantime this condition was used independently by Li [19] in finite dimensions for finite-valued differentiable convex functions. Moreover, in all the situations considered in the paper, the existence of Lagrange multipliers is obtained when, in addition, a certain overlapping of the domain of the objective function and of the admissible set holds.

The next section is devoted to the study of formulae for the normal cone and of related questions, under the weakest possible assumptions. As mentioned above, this plays an essential role in the existence of Lagrange multipliers. In the last section, we discuss the necessity of our hypotheses: metric regularity conditions, interiority conditions and the Slater's constraint qualification.

In fact, we realize a study of various constraint qualification conditions: conditions of type (BCQ), local metric regularity and the above metric regularity on bounded subsets (MRB), in general normed spaces and for extended-valued convex functions. Certain results along these lines have been established by the first author in the preprints [26, 27].

2. Constraint qualification conditions and formulae for the normal cone

Throughout this paper $(X, \|\cdot\|)$ is a real normed space whose topological dual X^* is endowed with the dual norm denoted also by $\|\cdot\|$; the closed unit balls of X and X^* are denoted by U_X and U_{X^*}, respectively. We denote by $\Lambda(X)$ the class of convex functions $h : X \to \mathbb{R} \cup \{\infty\}$ with nonempty domain $\operatorname{dom} h := \{x \in X \mid h(x) < \infty\}$, and by $\Gamma(X)$ the class of those functions $h \in \Lambda(X)$ which are also lower semicontinuous. Consider $h \in \Lambda(X)$ and assume that the (not necessarily closed) admissible set

$$C := [h \leq 0] := \{x \in X \mid h(x) \leq 0\} \tag{1}$$

is nonempty. Consider the minimization problem

$$\text{minimize} \quad g(x) \quad \text{s.t.} \quad h(x) \leq 0, \tag{2}$$

where $g \in \Lambda(X)$, too. Note that the problem (P) is equivalent with problem (2) by taking $h := \max_{i \in I} h_i$. In order that the problem (2) be non trivial we assume that $C \cap \operatorname{dom} g \neq \emptyset$. As mentioned in the Introduction, for deriving optimality criteria for the problem (2), a usual hypothesis is the Slater condition, *i.e.*,

$$\exists\, \widetilde{x} \in X \ : \ h(\widetilde{x}) < 0 \tag{3}$$

(which is equivalent to (SCQ) when $h := \max_{i \in I} h_i$). In fact, because we also envisage the case when g is not necessarily finite-valued, we slightly modify condition (3):

$$\exists\, \widetilde{x} \in \operatorname{dom} g \ : \ h(\widetilde{x}) < 0. \tag{4}$$

The next result is known for a long time when h is finite-valued (see [2, Th. 3.1.2], [3, Th. 3.1.2]). In the sequel we use the convention $0 \cdot \infty := \infty$.

Proposition 2.1. *Let $g, h \in \Lambda(X)$. Assume that condition* (4) *holds. Then $\overline{x} \in C \cap \operatorname{dom} g$ is a solution of problem* (2) *if and only if there exists $\lambda \geq 0$ such that $\lambda h(\overline{x}) = 0$ and $\overline{x}$ is a minimizer of $g + \lambda h$.*

Proof. Let $\overline{x} \in X$ be a minimizer of $g + \lambda h$ for some $\lambda \geq 0$ with $\lambda h(\overline{x}) = 0$. It is clear that $\overline{x} \in \operatorname{dom} g \cap \operatorname{dom} h$. If $x \in C$ then $g(\overline{x}) = (g+\lambda h)(\overline{x}) \leq (g+\lambda h)(x) = g(x) + \lambda h(x) \leq g(x)$, which proves that $\overline{x}$ is a solution of problem (2).

Conversely, assume that $\overline{x}$ is a solution of problem (2). Hence $\overline{x} \in C \cap \operatorname{dom} g$. Consider $F : X \times \mathbb{R} \to \overline{\mathbb{R}}$ defined by $F(x,t) := g(x)$ for $h(x) \leq t$, $F(x,t) := \infty$ otherwise. By hypothesis $(\widetilde{x}, 0) \in \operatorname{dom} F$ and $F(\widetilde{x}, \cdot)$ is continuous at 0. Using the stability theorem in Ekeland–Temam [13, Props. 2.2, 2.3 (ch. III)] or [31, Th. 2.7.1(iii)], there exists $\lambda \in \mathbb{R}$ such that

$$\begin{aligned} g(\overline{x}) &= \inf\nolimits_{x \in X} F(x,0) = \max\nolimits_{\mu \in \mathbb{R}} \big(-F^*(0,\mu) \big) = -F^*(0,-\lambda) \\ &= \inf\{ g(x) + \lambda\,(h(x) + s) \mid x \in \operatorname{dom} h,\ s \in \mathbb{R}_+ \}, \end{aligned}$$

where the conjugate $f^* : X^* \to \overline{\mathbb{R}}$ of the function $f : X \to \overline{\mathbb{R}}$ is defined by $f^*(x^*) := \sup_{x\in X} \left(\langle x, x^*\rangle - f(x)\right)$. It follows that $\lambda \geq 0$. Hence

$$g(\overline{x}) = \inf\{g(x) + \lambda h(x) \mid x \in \operatorname{dom} h\} \leq g(\overline{x}) + \lambda h(\overline{x}) \leq g(\overline{x}),$$

and so $\lambda h(\overline{x}) = 0$ and $\overline{x}$ is a minimizer of $g + \lambda h$. □

Note that if condition (3) holds but (4) does not, the conclusion of the preceding proposition may fail.

Example 2.2. Let $x^* \in X^* \setminus \{0\}$ and consider the functions $g, h : X \to \overline{\mathbb{R}}$ be defined by $h(x) := \langle x, x^*\rangle$, $g(x) := -\sqrt{\langle x, x^*\rangle}$ for $\langle x, x^*\rangle \geq 0$, $g(x) := \infty$ otherwise. It is obvious that condition (3) holds but (4) does not. Of course, 0 is a solution of problem (2). Assuming that 0 is a minimizer of $g + \lambda h$ for some $\lambda \geq 0$ we obtain that $0 \leq -\sqrt{t} + \lambda t$ for all $t \geq 0$, whence the contradiction $1 \leq \lambda\sqrt{t}$ for all $t > 0$.

Using the previous result one obtains the formula (5) below for the normal cone to C at $x \in C$ defined by

$$N(C, x) := \{x^* \in X^* \mid \langle y - x, x^*\rangle \leq 0,\ \forall\, y \in C\} = \partial\iota_C(x),$$

where ι_C is the indicator function of C given by $\iota_C(x) = 0$ for $x \in C$ and $\iota_C(x) = \infty$ for $x \in X \setminus C$; for the function $f \in \Lambda(X)$ the (Fenchel) subdifferential of f at $x \in \operatorname{dom} f$ is defined by

$$\partial f(x) := \{x^* \in X^* \mid \langle y - x, x^*\rangle \leq f(y) - f(x),\ \forall\, y \in X\}.$$

Proposition 2.3. *Let $h \in \Lambda(X)$ satisfy condition* (3)*. Consider $x \in C$, where C is defined by* (1)*. Then*

$$N(C, x) = \bigcup \left\{\partial(\lambda h)(x) \mid \lambda \geq 0,\ \lambda \cdot h(x) = 0\right\}, \tag{5}$$

where $0h$ means $\iota_{\operatorname{dom} h}$.

Proof. The inclusion "$\supset$" in (5) holds without any condition on h. Indeed, let $x^* \in \partial(\lambda h)(x)$ for some $\lambda \geq 0$ with $\lambda h(x) = 0$. Then for $y \in C$ we have that $\langle y - x, x^*\rangle \leq \lambda h(y) - \lambda h(x) = \lambda h(y) \leq 0$, and so $x^* \in N(C, x)$.

Let $x^* \in N(C, x)$. Then $\langle y - x, x^*\rangle \leq 0$ for $h(y) \leq 0$, *i.e.*, x is a solution of problem (2), where $g(y) := -\langle y, x^*\rangle$. Since $\operatorname{dom} g = X$, condition (4) holds. Applying Proposition 2.1 we get $\lambda \geq 0$ such that $\lambda h(x) = 0$ and x is a minimizer of $g + \lambda h$, *i.e.*, $-\langle x, x^*\rangle + \lambda h(x) \leq -\langle y, x^*\rangle + \lambda h(y)$ for every $y \in X$. Of course, this means that $x^* \in \partial(\lambda h)(x)$. □

When h is continuous at $x \in \operatorname{dom} h$ we have that $x \in \operatorname{int}(\operatorname{dom} h)$, and so $\partial(0h)(x) = N(\operatorname{dom} h, x) = \{0\}$. Moreover, if $h(x) = 0$ then relation (5) becomes the well-known formula for the normal cone: $N(C, x) = \mathbb{R}_+\partial h(x)$ (see Rockafellar [25, Cor. 23.7.1], Laurent [17, p. 388], Giles [14, p. 185]).

Proposition 2.3 was obtained by Penot and Zălinescu [22, Prop. 5.4] for $x \in [h = 0] := \{y \in X \mid h(y) = 0\}$. Also without assuming h to be continuous at $x \in [h = 0]$ (but still $\inf h < 0$), in finite dimensions Rockafellar [25, Th. 23.7] obtains that $N(C, x) = \operatorname{cl}\left(\mathbb{R}_+\partial h(x)\right)$ under the additional hypothesis that $\partial h(x) \neq \emptyset$, while Pshenichnyi [23, Th.

3.17] obtains that $N(C,x)=\mathbb{R}_+\partial h(x)$ under the additional hypothesis that the directional derivative $h'(x,\cdot)$ is lower semicontinuous, where

$$h'(x,u):=\lim_{t\to 0^+}\frac{h(x+tu)-h(x)}{t}. \tag{6}$$

Notice that one cannot replace the formula (5) by $N(C,x)=\mathbb{R}_+\partial h(x)$ when $h(x)=0$ and $x\notin \operatorname{core}(\operatorname{dom} h)$. Take for example $h(t)=-\sqrt{t}$ for $t\geq 0$, $h(t)=\infty$ for $t<0$ and $x=0$ in which case $\partial h(0)=\emptyset$. Note that Slater's condition (3) is not necessary for having formula (5). Indeed, consider $A\subset X$ a nonempty convex set and $x\in A$. Then $\operatorname{cl} A=\{y\in X \mid d_A(y)\leq 0\}$ and

$$N(A,x)=\mathbb{R}_+\partial d_A(x), \tag{7}$$

where $d_A(y):=\inf\{\|y-a\| \mid a\in A\}$ is the distance from $y\in X$ to A. Indeed, $d_A=\|\cdot\|\,\square\,\iota_A$ and the convolution is exact at $x\in A: d_A(x)=\|0\|+\iota_A(x)$. Hence, by a well-known formula (see Laurent [17, Prop. 6.6.4]), $\partial d_A(x)=\partial\|\cdot\|(0)\cap\partial\iota_A(x)=U_{X^*}\cap N(A,x)$. The relation (7) is now immediate.

When g is finite-valued and continuous, or more generally,

$$\exists\, x_0\in C\cap\operatorname{dom} g \;:\; g \text{ is continuous at } x_0, \tag{8}$$

an alternate proof of Proposition 2.1 is the following: $\overline{x}$ is a solution of (2) if and only if $\overline{x}$ is a minimizer of $g+\iota_C$ if and only if $0\in\partial(g+\iota_C)(\overline{x})=\partial g(\overline{x})+\partial\iota_C(\overline{x})$, the equality being true because g is continuous at some point of $C\cap\operatorname{dom} g$. So, using Proposition 2.3, $\overline{x}$ is a solution of (2) if and only if there exists $\lambda\geq 0$ such that $\lambda\cdot h(\overline{x})=0$ and $0\in\partial g(\overline{x})+\partial(\lambda h)(\overline{x})=\partial(g+\lambda h)(\overline{x})$.

In the above argument we may replace condition (8) by

$$\operatorname{int} C\cap\operatorname{dom} g\neq\emptyset. \tag{9}$$

The argument above shows that what is really needed for having the conclusion of Proposition 2.1 is formula (5) for the normal cone to C at $x\in C$. Hence we have proved the following result.

Proposition 2.4. *Let $g,h\in\Lambda(X)$ be such that condition (8) or (9) is verified. Assume that formula (5) holds for $x\in C\cap\operatorname{dom} g$. Then x is a solution of problem (2) if and only if there exists $\lambda\geq 0$ such that $\lambda h(x)=0$ and x is a minimizer of $g+\lambda h$.*

In fact, without having formula (5) for the normal cone, there is no hope for the conclusion of the preceding proposition to hold as the next result shows.

Proposition 2.5. *Assume that for any function $g\in\Lambda(X)$ satisfying (8) and any solution x_g of problem (2) there exists $\lambda_g\geq 0$ such that $\lambda_g h(x_g)=0$ and x_g is a minimizer of $g+\lambda_g h$. Then formula (5) holds for any $x\in C$.*

Proof. Let $x\in C$ and take $x^*\in N(C,x)$. Then x is a solution of problem (2) with $g:=-x^*$. Then, by hypothesis, there exists $\lambda\geq 0$ such that $\lambda h(x)=0$ and x is a minimizer of $-x^*+\lambda h$, which means that $x^*\in\partial(\lambda h)(x)$. Hence the inclusion "$\subset$" holds in (5). As the converse inclusion holds always (see the proof of Proposition 2.3), the conclusion follows. □

The proof above shows that in the preceding proposition one can replace $\Lambda(X)$ by X^*. Also note that combining Propositions 2.1 and 2.5 one obtains that the conclusion of Proposition 2.1 holds for any finite-valued convex function g if and only if formula (5) holds for every $x \in C$. So one generalizes Proposition VII.2.1.2 in [15] by taking into consideration the formula

$$\partial h(x) = \left\{ \sum\nolimits_{i\in I(x)} \lambda_i x_i^* \;\middle|\; \lambda_i \geq 0,\ x_i^* \in \partial h_i(x)\ \forall i \in I(x),\ \sum\nolimits_{i\in I(x)} \lambda_i = 1 \right\}, \tag{10}$$

where $h = \max_{i\in I} h_i$ with $h_i : X \to \overline{\mathbb{R}}$ proper convex functions, continuous at $x \in \bigcap_{i\in I} \operatorname{dom} h_i$ and $I(x) := \{i \in I \mid h_i(x) = h(x)\}$ (see Tikhomirov [29, Th. 7]). In fact a similar formula holds even if h_i are not continuous at some point, as can be seen in [31, Cor. 2.8.11]. In particular such a formula applies to the function $h_+ := \max\{h, 0\}$. In this case, for $h \in \Lambda(X)$, we get (see [31, Example 2.8.1])

$$\partial h_+(x) = \begin{cases} \partial h(x) & \text{if } \ h(x) > 0, \\ \bigcup \{\partial(\lambda h)(x) \mid \lambda \in [0,1]\} & \text{if } \ h(x) = 0, \\ \partial(0h)(x) = \partial\iota_{\operatorname{dom} h}(x) & \text{if } \ h(x) < 0. \end{cases} \tag{11}$$

In the case $h = \max_{i\in I} h_i$ with $h_i : \mathbb{R}^n \to \mathbb{R}$ differentiable convex functions and x satisfies $h(x) = 0$ formula (5) becomes (ACQ') mentioned in the Introduction.

Another sufficient condition for the validity of formula (5) is given in the next result.

Proposition 2.6. *Let $h \in \Lambda(X)$ and $x \in [h = 0]$. Then (5) holds provided that*

$$\exists\, r, \gamma > 0,\ \forall\, y \in x + rU_X \ :\ h_+(y) \geq \gamma \cdot d_C(y). \tag{12}$$

Proof. Consider $x^* \in N(C, x)$. From formula (7) we have that $x^* = \mu u^*$ for some $\mu \geq 0$ and $u^* \in \partial d_C(x)$. It follows that $\gamma u^* \in \partial(\gamma d_C)(x)$. Since, by hypothesis, $\gamma d_C \leq h_+ + \iota_{x+rU_X}$ and both functions coincide at x, we obtain that $\gamma u^* \in \partial(h_+ + \iota_{x+rU_X})(x) = \partial h_+(x) + \partial\iota_{x+rU_X}(x) = \partial h_+(x)$, the first equality being true since ι_{x+rU_X} is finite and continuous at x. Hence $x^* \in \mathbb{R}_+\partial h_+(x)$ which together with (11) shows that the inclusion $\subset$ holds in (5). The converse inclusion being always true, the conclusion follows. □

When $h \in \Gamma(\mathbb{R}^n)$ and $x \in \operatorname{int}(\operatorname{dom} h)$ the preceding result can be deduced from Lewis and Pang [20, Prop. 2]. When (12) holds at $x \in C$, Li [19] says that the system $h(y) \leq 0$, $y \in X$, is metrically regular at x, and the preceding system is metrically regular when (12) holds at any $x \in C$ (see also Deng [11]).

If $C_1, \dots, C_n \subset X$ are closed convex sets and $x \in C := C_1 \cap \dots \cap C_n$, one says (see Pang [21]) that $\{C_1, \dots, C_n\}$ is metrically regular at x if there exists $\gamma, r > 0$ such that

$$\max\{d_{C_i}(y) \mid 1 \leq i \leq n\} \geq \gamma \cdot d_C(y) \quad \forall\, y \in x + rU_X.$$

Taking $h := \max\{d_{C_i} \mid 1 \leq i \leq n\}$, we have that $C = [h \leq 0]$, and the above relation means that the system $h(y) \leq 0$ is metrically regular at x. Hence, by the preceding proposition and formulae (7) and (10) we get

$$N(C, x) = N(C_1, x) + \dots + N(C_n, x) \tag{13}$$

when C is metrically regular at $x \in C$. So we recover Proposition 6 of Pang [21]. Note that Bauschke, Borwein and Li [6] say that $\{C_1, \dots, C_n\}$ satisfies the "strong CHIP" condition when (13) holds for every $x \in C$.

Taking into account that, by Proposition 2.3, formula (5) holds when $h(x) < 0$ (because condition (3) holds in this case), the next result follows immediately.

Corollary 2.7. *Let $h \in \Lambda(X)$. If*

$$\forall\, x \in [h = 0],\ \exists\, r_x, \gamma_x > 0,\ \forall\, y \in x + r_x U_X\ :\ h_+(y) \geq \gamma_x \cdot d_C(y), \tag{14}$$

then formula (5) *holds for every $x \in C$.*

In the next result we show that Slater's constraint qualification (3) is strictly stronger than the condition

$$\forall\, r > 0,\ \exists\, \gamma_r > 0,\ \forall\, x \in rU_X\ :\ h_+(x) \geq \gamma_r \cdot d_C(x); \tag{15}$$

when $h = \max_{i \in I} h_i$, condition (15) is nothing else but condition (MRB) from the Introduction. When $h = \max\{d_{C_1}, \dots, d_{C_n}\}$ condition (15) means that $\{C_1, \dots, C_n\}$ is "boundedly linearly regular" in the sense of Bauschke and Borwein [5] (see also [4]).

Proposition 2.8. *Let $h \in \Lambda(X)$. Assume that $h(\widehat{x}) < 0$. Then*

$$\forall\, r > 0,\ \forall\, x \in \widehat{x} + rU_X\ :\ h_+(x) \geq -r^{-1} h(\widehat{x}) \cdot d_C(x). \tag{16}$$

Proof. With our hypothesis, Robinson in [24] proved that

$$d_C(x) \leq \big(- h(\widehat{x})\big)^{-1} \|x - \widehat{x}\| \cdot h_+(x) \quad \forall\, x \in X, \tag{17}$$

whence (16) follows immediately. □

A simple direct proof of (16) can be found in [10, Example 4.2] (see also [18], [26], [27] and [19]).

An alternative proof for the fact that formula (5) holds for every $x \in C$ when condition (15) does is obtained using the next result, result which is also interesting for itself. Here the multi-valued operator $F_X : X \rightrightarrows X^*$ is the duality mapping of X.

Proposition 2.9. *Let X be a reflexive Banach space and $h \in \Gamma(X)$. If condition* (15) *holds then for all $\lambda > 0$ and $y \in X$ there exists $x_{\lambda,y} \in X$ such that*

$$0 \in F_X(x_{\lambda,y} - y) + \lambda \partial h(x_{\lambda,y}). \tag{18}$$

Moreover, for every $y \in X$ there exists $\lambda_y > 0$ such that $x_{\lambda,y} \in C$ for $\lambda > \lambda_y$.

Proof. The function $\varphi_{\lambda,y} : X \to \overline{\mathbb{R}}$, $\varphi_{\lambda,y}(x) := \frac{1}{2}\|x - y\|^2 + \lambda h(x)$ is lower semicontinuous, convex and coercive. Because X is reflexive, there exists $x_{\lambda,y} \in X$ such that $\varphi_{\lambda,y}(x_{\lambda,y}) \leq \varphi_{\lambda,y}(x)$ for every $x \in X$, and so $0 \in \partial\varphi_{\lambda,y}(x_{\lambda,y})$, which is equivalent to (18).

For the second part we consider first the case when the infimum of h is attained. Hence there exists $\overline{x} \in \operatorname{dom} h$ such that $0 \in \partial h(\overline{x})$. We have that

$$\|x_{\lambda,y} - y\|^2 + 2\lambda h(x_{\lambda,y}) \leq \|\overline{x} - y\|^2 + 2\lambda h(\overline{x}) \leq \|\overline{x} - y\|^2 + 2\lambda h(x_{\lambda,y}),$$

whence $\|x_{\lambda,y} - y\| \leq \|\overline{x} - y\|$. It follows that $\|x_{\lambda,y}\| \leq \|y\| + \|\overline{x} - y\|$. Let $y \in Y$ be fixed. Consider $r > \|y\| + \|\overline{x} - y\|$; take $\gamma := \gamma_r$ with $\gamma_r > 0$ given by (15) and denote $x_{\lambda,y}$ by x_λ. Assume that $x_\lambda \notin C$. Then there exists $x'_\lambda \in P_C(x_\lambda)$. Taking into account (15), we have that

$$2\lambda\gamma d_C(x_\lambda) + \|x_\lambda - y\|^2 \leq 2\lambda h(x_\lambda) + \|x_\lambda - y\|^2 \leq 2\lambda h(x'_\lambda) + \|x'_\lambda - y\|^2 \leq \|x'_\lambda - y\|^2 ,$$

and so

$$2\lambda\gamma \|x_\lambda - x'_\lambda\| \leq \|x'_\lambda - y\|^2 - \|x_\lambda - y\|^2 \leq \|x_\lambda - x'_\lambda\| \left(\|x'_\lambda - y\| + \|x_\lambda - y\|\right).$$

Hence

$$\begin{aligned} 2\lambda\gamma &\leq \|x'_\lambda - y\| + \|x_\lambda - y\| \leq 2\|x_\lambda - y\| + \|x_\lambda - x'_\lambda\| \leq 2\|x_\lambda - y\| + \|x_\lambda - \overline{x}\| \\ &\leq 3\|x_\lambda - y\| + \|y - \overline{x}\| \leq 4\|\overline{x} - y\| . \end{aligned}$$

Thus the conclusion holds for $\lambda_y := 2\gamma^{-1}\|\overline{x} - y\|$. Assume now that h does not attain its infimum. Of course, condition (15) is satisfied by h_+ and h_+ attains its infimum. For $y \in X$ and $\lambda > 0$ denote by $x^+_{\lambda,y}$ the element $x \in X$ satisfying $0 \in F_X(x - y) + \lambda\partial h_+(x)$. It is obvious that $\frac{1}{2}\|x - y\|^2 + \lambda h(x) \leq \frac{1}{2}\|x - y\|^2 + \lambda h_+(x)$ for every $x \in X$. Let $\lambda_y > 0$ be such that $x^+_{\lambda,y} \in C$ for $\lambda > \lambda_y$. Take $\lambda > \lambda_y$ and assume that $x_{\lambda,y} \notin C$. Then

$$\begin{aligned} \|x_{\lambda,y} - y\|^2 + 2\lambda h(x_{\lambda,y}) &\leq \|x^+_{\lambda,y} - y\|^2 + 2\lambda h_+(x^+_{\lambda,y}) \leq \|x_{\lambda,y} - y\|^2 + 2\lambda h_+(x_{\lambda,y}) \\ &= \|x_{\lambda,y} - y\|^2 + 2\lambda h(x_{\lambda,y}). \end{aligned}$$

It follows that $x_{\lambda,y}$ minimizes $\frac{1}{2}\|\cdot - y\|^2 + \lambda h_+(\cdot)$, and so, by our choice of λ_y, $x_{\lambda,y} \in C$. This contradiction proves our assertion. □

Proposition 2.10. *Let X be a reflexive Banach space and $h : X \to \mathbb{R}$ be convex and continuous, satisfying condition* (15). *Then the multi-valued operator $N : X \rightrightarrows X^*$ defined by*

$$N(x) = \begin{cases} \{0\} & \text{if } h(x) < 0, \\ \mathbb{R}_+\partial h(x) & \text{if } h(x) = 0, \\ \emptyset & \text{if } h(x) > 0, \end{cases} \tag{19}$$

is maximal monotone and $N(x) = \partial\iota_C(x) = N(C, x)$ for every $x \in C$.

Note that the relation $N(x) = N(C, x)$ for every $x \in C$ (and so $N(x) = \partial\iota_C(x)$ for every $x \in X$) follows from Corollary 2.7 because (15) ⇒ (14) and $C \subset \operatorname{int}(\operatorname{dom} h) = X$. Moreover, because ι_C is lower semicontinuous and X is a Banach space the maximality of N follows from Rockafellar's theorem. However we give a direct proof using the preceding result for the reader who is more familiar with monotone operators.

Proof. Note first that $N = N_+$, where N_+ denotes the operator defined by (19) which corresponds to h_+; just use formula (11) and take into account that $0h = 0$ in our conditions. So, we assume that $h \geq 0$.

The space X being reflexive, by a classic renorming theorem (see Diestel [12, Cor. 2, p. 148]) we may assume that X is strictly convex and smooth, and so F_X is single-valued. By the definition of the subdifferential, it is clear that $\operatorname{gr} N \subset \operatorname{gr} \partial\iota_C$ and, therefore, that

N is monotone. We show that N is maximal monotone in $X \times X^*$ and this will give the desired equality. In order to apply the converse part in Minty's theorem (see [1]), it is sufficient to show that the equation

$$F_X(x-y) + N(x) \ni 0$$

has solutions for every fixed $y \in X$. Let $y \in X$ be fixed and consider $\lambda > \lambda_y$, where λ_y is given by the preceding proposition. Then the solution $x_{\lambda,y}$ of equation (18) is in C, and so $x_{\lambda,y}$ is a solution of the equation displayed above. Therefore N is maximal monotone, and so $N = \partial\iota_C$. □

At the end of this section we give an application to abstract control problems:

$$\min\ L(y,u), \tag{20}$$

subject to

$$Ay = Bu + \varphi, \tag{21}$$

$$h(y,u) \le 0, \tag{22}$$

where U, Y are Hilbert spaces with $Y \subset X \subset Y^*$, $L, h : Y \times U \to \mathbb{R}$ are convex continuous mappings, $\varphi \in Y^*$ and $A : Y \to Y^*$, $B : U \to Y^*$ are linear bounded operators with

$$\langle Ay, y\rangle_{Y^*\times Y} \ge \omega|y|_Y^2 \quad \forall\, y \in Y, \tag{23}$$

for some $\omega > 0$.

The set $C = \{(y,u) \in Y \times U \mid h(y,u) \le 0\}$ is a closed convex set and we assume that there is an admissible pair $(\widetilde{y}, \widetilde{u})$ such that $(\widetilde{y}, \widetilde{u}) \in \operatorname{int} C$ (w.r.t. the norm topology of $Y \times U$). Typical situations of problem (20)–(22) are obtained when Y, Y^* are Sobolev spaces, X, U are Lebesgue spaces and A is an (elliptic) partial differential operator, while B is some distributed or boundary control action.

Since condition (23) holds, the equation (21) has a unique solution for any $u \in U$. By shifting the domains of L, h and redenoting the obtained mappings again by L, h, we may assume $\varphi = 0$. We also notice the generality of the mixed constraint (22) which includes both state and control constraints.

We shall apply the previous results. We consider the closed subspace $K = \{(y,u) \in Y \times U \mid Ay = Bu\}$ and we replace L by $L + \iota_K$ in (20). If h satisfies (15) and $(\overline{y}, \overline{u})$ is an optimal pair for (20)–(22), then Propositions 2.6 and 2.4 show that there is $\lambda \ge 0$ such that

$$0 \in \partial L(\overline{y}, \overline{u}) + \partial\iota_K(\overline{y}, \overline{u}) + \lambda\partial h(\overline{y}, \overline{u}) \tag{24}$$

and $\lambda h(\overline{y}, \overline{u}) = 0$. Here, we also use that $K \cap \operatorname{int} C \ne \emptyset$ in order to apply the additivity rule for the subdifferential. It is known that $\partial\iota_K(\overline{y}, \overline{u}) = K^\perp$ and a simple calculus (taking into account that A^* is an isomorphism under condition (23)) shows that

$$K^\perp = \{(A^*p, -B^*p) \mid p \in Y\} \subset Y^* \times U^*. \tag{25}$$

By (24), (25) we infer the optimality conditions for the problem (20)–(22):

$$\begin{aligned} -A^*\overline{p} &\in \partial_1 L(\overline{y},\overline{u}) + \lambda\partial_1 h(\overline{y},\overline{u}),\\ B^*\overline{p} &\in \partial_2 L(\overline{y},\overline{u}) + \lambda\partial_2 h(\overline{y},\overline{u}),\\ &\lambda h(\overline{y},\overline{u}) = 0, \quad \lambda \geq 0, \end{aligned}$$

where $\partial_i L, \partial_i h$, $i = 1, 2$, denote the i-th component of the ordered pairs $\partial L, \partial h$ and not a partial subdifferential.

In the work of Tiba and Bergounioux [28], a weaker form of the optimality system is obtained, without imposing interiority assumptions on C.

3. Necessary conditions

One can ask what constraint qualification conditions are necessary for the existence of Lagrange multipliers. This problem is discussed in this section.

3.1. Necessity of metric regularity conditions for the existence of Lagrange multipliers

Taking into account Propositions 2.4 and 2.5, the question raised above can be rephrased as follows: is the metric regularity condition (12) necessary for having formula (5)? A partial answer is given in the next result, where $h'(x,u)$ is defined by (6).

Proposition 3.1. *Let X be finite dimensional and $h \in \Gamma(X)$ be continuous at $x \in [h = 0]$. Assume that $h'(x_n,u_n) \to h'(x,u)$ for all sequences $(x_n) \subset C$ converging to x and all sequences (u_n) converging to u with $u_n \in F_X^{-1}(N(C,x_n))$ for every $n \in \mathbb{N}$. If formula* (5) *holds then condition* (12) *holds, too.*

Proof. Assume that formula (5) holds but condition (12) doesn't. Then there exist $(x_n) \subset X$ converging to x and $(\gamma_n) \subset]0,\infty[$ converging to 0 such that $0 < h(x_n) < \gamma_n d_C(x_n) \leq \gamma_n \|x_n - x\|$ for every $n \in \mathbb{N}$. The set C being a nonempty closed and convex subset of a finite dimensional normed space, there exists $\overline{x}_n \in P_C(x_n)$, where $P_A(y) := \{a \in A \mid d_A(y) = \|y - a\|\}$. It follows that $(\overline{x}_n)$ converges to x. Moreover, because $h|_{[x_n,\overline{x}_n]}$ is continuous, we have that $h(\overline{x}_n) = 0$, and so

$$h'(\overline{x}_n, x_n - \overline{x}_n) \leq h(x_n) - h(\overline{x}_n) < \gamma_n \|x_n - \overline{x}_n\| \quad \forall n \in \mathbb{N}. \tag{26}$$

Let $u_n := \|x_n - \overline{x}_n\|^{-1}(x_n - \overline{x}_n)$. By a known characterization of best approximations (see for example Th. 3.8.4(iv) in [31]), we have that $F_X(x_n - \overline{x}_n) \cap N(C,\overline{x}_n) \neq \emptyset$, and so $u_n \in F_X^{-1}(N(C,x_n))$ for every $n \in \mathbb{N}$. Consider $u_n^* \in F_X(u_n) \cap N(C,\overline{x}_n)$. Since X is finite dimensional we may assume that (u_n) converges to some $u \in X$ and (u_n^*) converges to $u^* \in X^*$. It follows that $u^* \in F_X(u) \cap N(C,x)$. Since formula (5) holds, there exists $\mu \geq 0$ such that $u^* \in \partial(\mu h)(x)$. Because $x \in \operatorname{int}(\operatorname{dom} h)$, h being continuous at x, and $u^* \neq 0$ we have that $\mu > 0$. Hence $u^* = \mu x^*$ for some $x^* \in \partial h(x)$. Then, by our hypothesis and (26), we obtain that $h'(x,u) \leq 0$, and so we get the contradiction $\mu = \langle u, x^*\rangle \leq h'(x,u) \leq 0$. The proof is complete. □

Note that when h is continuous and Gâteaux differentiable at $x \in \operatorname{dom} h$ then $h'(x_n,u_n) \to h'(x,u)$ for all sequences $(x_n) \subset \operatorname{dom} h$ converging to x and all sequences $(u_n) \subset X$

converging to u (see [9, Fact 2.3], [31, Th. 3.3.2]). Also note that the condition on h in the hypothesis of the preceding result is verified when h is finite and continuous on X and $\big(h'(x_n, u_n)\big) \to h'(x, u)$ whenever $(x_n) \to x$, $(u_n) \to u$ and $x_n + \lambda u_n \in \mathrm{cl}(X \setminus C)$ for all $\lambda \geq 0$.

Related to the hypothesis of Proposition 3.1, note that Bauschke, Borwein and Tseng [7, Cor. 3.2] gave an example of a closed convex cone $K \subset \mathbb{R}^4$ such that $\{K, Y\}$, with $Y := \mathbb{R}^3 \times \{0\}$, satisfies the strong CHIP condition but is not boundedly linearly regular.

When h is continuous at every point of $[h < 0] := \{y \in X \mid h(y) < 0\}$ it is obvious that (14) is equivalent with

$$\forall\, x \in C,\ \exists\, r_x, \gamma_x > 0,\ \forall\, y \in x + r_x U_X \ :\ h_+(y) \geq \gamma_x \cdot d_C(y), \tag{27}$$

but we do not know if this is true for general h (excepting the case $X = \mathbb{R}$). A condition stronger than condition (27) is condition (15). Condition (15) is intermediate between condition (27) and the existence of a *global error bound* for the (convex) inequality system $h(y) \leq 0$:

$$\exists\, \gamma > 0,\ \forall\, x \in X \ :\ h_+(x) \geq \gamma \cdot d_C(x). \tag{28}$$

It is obvious that (28) $\Rightarrow$ (15) $\Rightarrow$ (27) $\Rightarrow$ (14). It is known that Slater's condition (3) does not imply condition (28); see [16] for an example.

In the next result we point out a situation when conditions (15) and (27) are equivalent.

Proposition 3.2. *Let X be finite dimensional and $h \in \Gamma(X)$. Then the conditions* (15) *and* (27) *are equivalent.*

Proof. Assume that (27) holds but (15) does not. Then for some $r > 0$ and every $n \in \mathbb{N}$ there exists $x_n \in rU_X$ such that $0 < h(x_n) < \gamma_n \cdot d_C(x_n)$, where $(\gamma_n) \subset]0, \infty[$ is a sequence converging to 0. Since (x_n) is bounded we may assume that (x_n) converges to some $\overline{x} \in X$. It follows that $h(\overline{x}) \leq \liminf h(x_n) \leq 0 \cdot d_C(\overline{x}) = 0$. Hence $\overline{x} \in C$. By hypothesis there exist $\overline{r}, \overline{\gamma} > 0$ such that $h_+(x) \geq \overline{\gamma} \cdot d_C(x)$ for $x \in \overline{x} + \overline{r}U_X$. Taking n sufficiently large in order that $\|x_n - \overline{x}\| < \overline{r}$ and $\gamma_n < \overline{\gamma}$, we get a contradiction. □

Note that Li [19], as recalled above, considered the case when $h = \max_{i \in I} h_i$ with $h_i : \mathbb{R}^n \to \mathbb{R}$ differentiable convex functions. In this case he proved Proposition 3.2 as well as the equivalence of metric regularity and the validity of Abadie's CQ (both on C).

Remark 3.3. Imposing the (norm) compactness of C in the preceding proposition one gets (28) in any normed vector space X; for this use [10, Cor. 5.3] or [31, Cor. 3.4.2] and a classic argument with open covers. In the absence of this condition, in general reflexive Banach spaces the result is still open.

It is known (see Lewis and Pang [20]) that in the case $X = \mathbb{R}^n$ endowed with the Euclidean norm, condition (28) holds if and only if

$$\exists\, \gamma > 0,\ \forall\, x \in [h = 0],\ \forall\, u \in N(C, x) \ :\ h'(x, u) \geq \gamma \cdot \|u\| .$$

In the case of arbitrary normed spaces the preceding condition must be written as

$$\exists\, \gamma > 0,\ \forall\, x \in [h = 0],\ \forall\, u \in F_X^{-1}\left(N(C, x)\right) \ :\ h'(x, u) \geq \gamma \cdot \|u\| . \tag{29}$$

In [30] it is shown that (28) and (29) are equivalent (with the same γ). One can ask if there are similar characterizations for (15), (12) and (27). The answer is given in the next proposition.

Proposition 3.4. *Assume that X is a reflexive Banach space and $h \in \Gamma(X)$. Then condition* (15) *is equivalent to any one of the conditions*

$$\forall r > 0,\ \exists \gamma_r > 0,\ \forall x \in rU_X \cap [h = 0],\ \forall u \in F_X^{-1}(N(C, x)) \ :\ h'(x, u) \geq \gamma_r \cdot \|u\|, \tag{30}$$

$$\forall r > 0,\ \exists \gamma_r > 0,\ \forall x \in rU_X \cap C,\ \forall u \in F_X^{-1}(N(C, x)) \ :\ h'(x, u) \geq \gamma_r \cdot \|u\|, \tag{31}$$

and condition (12)*, for a fixed $x \in C$, is equivalent to*

$$\exists \delta, \gamma > 0,\ \forall y \in (x + \delta U_X) \cap C,\ \forall u \in F_X^{-1}(N(C, y)) \ :\ h'(y, u) \geq \gamma \cdot \|u\|. \tag{32}$$

Proof. (31) $\Rightarrow$ (30) is obvious.

(15) $\Rightarrow$ (31) Let $r > 0$ and take $r' > r$. Let $\gamma_r := \gamma_{r'} > 0$. Consider $x \in rU_X \cap C$ and $u \in F^{-1}(N(C, x))$. If $u = 0$ there is nothing to prove, so let $u \neq 0$. It is well known that $x \in P_C(x + tu)$ for every $t > 0$. In particular $x + tu \notin C$ for $t > 0$ because $d(x + tu, C) = t\|u\| > 0$. Take $t' = (r' - r)/\|u\|$; then $x + tu \in r'U_X$ for every $t \in [0, t']$. If $x + tu \notin \operatorname{dom} h$ for any $t > 0$ then $h'(x, u) = \infty$, and so $h'(x, u) \geq \gamma_r \|u\|$. Assume that $x + t''u \in \operatorname{dom} h$ for some $t'' > 0$; it follows that $h(x) = 0$. Otherwise, as $h|_{[x, x+t''u]}$ is continuous, there exists $\theta \in]0, 1[$ such that $h(x + \theta t''u) = 0$, contradicting the fact $d(x + tu, C) = t\|u\|$. As $x + tu \in r'U_X \setminus C$ for $0 < t \leq t_0 := \min\{t', t''\}$, from (15) we obtain that

$$h(x + tu) - h(x) \geq \gamma_r d(x + tu, C) = \gamma_r t\|u\| \quad \forall t \in]0, t_0],$$

and so $h'(x, u) \geq \gamma_r \|u\|$.

(30) $\Rightarrow$ (15) Let $c \in C$ be a fixed element, $r > 0$ and take $r' := 2r + \|c\|$. Take $\gamma_r := \gamma_{r'}$. Consider $y \in rU_X \setminus C$. If $y \notin \operatorname{dom} h$, the inequality $h(y) \geq \gamma_r d(y, C)$ is obvious; assume that $y \in \operatorname{dom} h$. Since X is a reflexive Banach space and C is closed and convex, there exists $x \in P_C(y)$. Assume that $h(x) < 0$. Because $h|_{[y,x]}$ is continuous, there exists $z \in]x, y[$ such that $h(z) = 0$. So we obtain the contradiction $\|x - y\| = d(y, C) \leq \|y - z\| < \|x - y\|$. Therefore $x \in [h = 0]$. Moreover, $\|x\| \leq \|x - y\| + \|y\| \leq \|y - c\| + \|y\| \leq 2r + \|c\| = r'$. From the characterization of the best approximations we have that $F_X(y - x) \cap N(C, x) \neq \emptyset$, and so $y - x \in F_X^{-1}(N(C, x))$. From our hypothesis we obtain that

$$h(y) = h(y) - h(x) \geq h'(x, y - x) \geq \gamma_{r'} \|y - x\| = \gamma_r \cdot d(y, C),$$

and so (15) holds.

(12) $\Rightarrow$ (32) Let $x \in C$ and $r, \gamma > 0$ be given by (12). Consider $\delta := r/2$. Let $y \in (x + \delta U_X) \cap C$ and $u \in F^{-1}(N(C, y)) \setminus \{0\}$. We have either $y + tu \notin \operatorname{dom} h$ for every $t > 0$ (and so $h'(y, u) = \infty$) or $y + t''u \in \operatorname{dom} h$ for some $t'' > 0$. In the last case, as in the proof of (15) $\Rightarrow$ (31), $h(y) = 0$ and for $t > 0$ sufficiently small we have that $y + tu \in (x + rU_X) \setminus C$. The conclusion follows similarly.

(32) $\Rightarrow$ (12) Let $x \in C$ be a fixed element and take $\delta, \gamma > 0$ given by (32). Consider $r := \delta/2$. Let $y \in (x + \delta U_X) \setminus C$. We may take $y \in \operatorname{dom} h$ and consider $y' \in P_C(y)$.

As above we obtain that $y' \in [h = 0] \subset C$. Moreover, $\|y' - x\| \leq \|y' - y\| + \|y - x\| \leq 2\|x - y\| \leq r$. Hence, as in the proof of (30) $\Rightarrow$ (15), $y - y' \in F_X^{-1}(N(C, y'))$. From our hypothesis we obtain that

$$h(y) = h(y) - h(y') \geq h'(y', y - y') \geq \gamma \|y - y'\| = \gamma \cdot d(y, C),$$

and so (12) holds. The proof is complete. □

The preceding result furnishes a characterization of the metric regularity of the intersection of a finite number of convex sets as an answer to the following remark of Pang [21, p. 314]: "there is no known characterization for the local metric regularity of these convex sets C_i at a point in their intersection".

Corollary 3.5. *Let X be a reflexive Banach space and C_i, $1 \leq i \leq n$, be closed convex subsets of X. Then $C := \bigcap_{1 \leq i \leq n} C_i$ is metrically regular at $x \in C$ if and only if*

$$\exists \gamma, r > 0,\ \forall y \in C \cap (x + rU_X),\ \forall u \in F_X^{-1}(N(C, y))\ :\ \max_{1 \leq i \leq n} d_{\mathcal{C}(C_i, x)}(u) \geq \gamma \cdot \|u\|,$$

where $\mathcal{C}(C_i, x) := \operatorname{cl}\big(\operatorname{cone}(C_i - x)\big)$.

Proof. It is sufficient to observe that for $h = \max_{i \in I} h_i$ with $h_i \in \Lambda(X)$, and $x \in \operatorname{core}(\operatorname{dom} h)$ we have that

$$h'(x, u) = \max\nolimits_{i \in I(x)} h_i'(x, u) \quad \forall u \in X,$$

where $I(x) := \{i \in I \mid h_i(x) = h(x)\}$, and that for a nonempty closed convex set $A \subset X$ and $x \in A$, $d_A'(x, u) = d(u, \mathcal{C}(A, x))$ (see Lewis and Pang [20] for the case of Euclidean spaces and Zălinescu [30, 31] for the general case). Applying the preceding proposition the conclusion follows. □

Summarizing the preceding results we have the following implications, where MRB means condition (15) (*i.e.*, metric regularity on bounded sets), MR means condition (27) (*i.e.*, metric regularity), MRx means condition (12) (*i.e.*, metric regularity at x) and $C(X)$ means the class of all continuous convex functions from X to $\mathbb{R}$.

$$\forall g \in C(X),\ \exists \text{ Lagrange multiplier } \lambda \overset{\text{Props. 2.4, 2.5}}{\Longleftrightarrow} \text{(5) holds } \forall x \in C;$$

$$\text{Slater's CQ (3)} \overset{\text{Prop. 2.8}}{\Longrightarrow} \text{MRB} \overset{\text{obvious}}{\Longrightarrow} \text{MR} \overset{\text{obvious}}{\Longleftrightarrow} \text{MR}x \text{ holds } \forall x \in C;$$

$$\text{MR}x \overset{\text{Cor. 2.7}}{\Longrightarrow} \text{(5)}, \quad \forall x \in C.$$

If $\dim X < \infty$ and $h \in \Gamma(X)$:

$$\text{MRB} \overset{\text{Prop. 3.2}}{\Longleftarrow} \text{MR; MR}x \overset{\text{Prop. 3.1 (+ suppl. conditions)}}{\Longleftarrow} \text{(5)}.$$

The results above show that Slater's condition is too strong for convex optimization. What is really needed is formula (5) for the normal cone to C, and this formula is guaranteed by a metric regularity condition.

3.2. Necessity of other assumptions

As mentioned in the preceding section, Slater's condition (4) is too strong for convex optimization. On the other hand, Example 1 shows that Slater's CQ (3) does not ensure the existence of Lagrange multipliers when neither condition (8) nor condition (9) is satisfied. The next result shows that for every lower semicontinuous function h (convex or not) satisfying very weak differentiability hypotheses, there exists a function $g \in \Gamma(X)$ such that no Lagrange multipliers exist.

Proposition 3.6. *Assume that the proper and lower semicontinuous function $h : X \to \overline{\mathbb{R}}$ is such that $\emptyset \neq C := [h \leq 0] \neq X$ and $\overline{D}h(\overline{x}, u) < \infty$ for all $\overline{x} \in C$ and any $u \in X$ with $\overline{x} + tu \notin C$ for $t \in]0, t_u]$ (for some $t_u > 0$), where*

$$\overline{D}h(\overline{x}, u) := \limsup_{t \to 0+} \frac{h(\overline{x} + tu) - h(\overline{x})}{t}.$$

Then there exists $g \in \Gamma(X)$ and a (local) minimizer $\overline{x} \in C$ of g on C such that $\overline{x}$ is not a local minimizer of $g + \eta h$ on X for any $\eta \geq 0$ with $\eta h(\overline{x}) = 0$.

Proof. There exist $\overline{x} \in C$ and $\widehat{x} \in X \setminus C$ such that $[\widehat{x}, \overline{x}[\subset X \setminus C$. Indeed, taking $\widetilde{x} \in C$ and $\overline{\lambda} := \inf\{\lambda \in [0,1] \mid (1-\lambda)\widehat{x} + \lambda\widetilde{x} \in C\}$ we have that $\overline{\lambda} > 0$ and $\overline{x} := (1-\overline{\lambda})\widehat{x} + \overline{\lambda}\widetilde{x}$ has the desired property because C is closed. Consider

$$g : X \to \overline{\mathbb{R}}, \quad g(x) := \begin{cases} -\sqrt{t} & \text{if} \quad x = (1-t)\overline{x} + t\widehat{x} \text{ with } t \in [0,1], \\ \infty & \quad\quad\ \text{otherwise.} \end{cases}$$

It is obvious that $g \in \Gamma(X)$ and $C \cap \operatorname{dom} g = \{\overline{x}\}$. Therefore $\overline{x}$ is a (local) minimizer of g on C. Assume that $\eta \geq 0$ is such that $\eta h(\overline{x}) = 0$ and $\overline{x}$ is a local minimizer of $g + \eta h$. Because $\overline{D}h(\overline{x}, \widehat{x} - \overline{x}) < \infty$ we have that $(1-t)\overline{x} + t\widehat{x} \in \operatorname{dom} h$ for $t \in]0, t_0]$, for some $t_0 \in]0, 1]$. Then there exists $t_1 \in]0, t_0]$ such that

$$0 = g(\overline{x}) + \eta h(\overline{x}) \leq -\sqrt{t} + \eta h(\overline{x} + t(\widehat{x} - \overline{x}))$$

for all $t \in [0, t_1]$. It follows that $\eta > 0$ and

$$\eta \frac{h(\overline{x} + t(\widehat{x} - \overline{x})) - h(\overline{x})}{t} \geq \frac{1}{\sqrt{t}}$$

for $t \in]0, t_1]$. Taking the lim sup for $t \to 0+$, we get the contradiction $\eta\overline{D}h(\overline{x}, \widehat{x} - \overline{x}) = \infty$. □

Note that we may ask only that C be closed instead of assuming h to be lower semicontinuous in the preceding result.

Another situation when Lagrange multipliers do not exist is furnished in the next result.

Proposition 3.7. *Assume that the proper function $h : X \to \overline{\mathbb{R}}$ is such that $\emptyset \neq C := [h \leq 0] \subset \operatorname{core}(\operatorname{dom} h)$ and $\underline{D}h(x, u) \leq 0$ for all $x \in C$ and any $u \in X$, where $\operatorname{core}(\operatorname{dom} h)$ is the algebraic interior of $\operatorname{dom} h$ and*

$$\underline{D}h(\overline{x}, u) := \liminf_{t \to 0+} \frac{h(\overline{x} + tu) - h(\overline{x})}{t}.$$

If $g \in \Lambda(X)$ and $\overline{x} \in C \cap \operatorname{dom} g$ are such that $g(\overline{x}) > \inf g$ then there are no Lagrange multipliers for (P) at $\overline{x}$.

Proof. Assume that $\overline{x} \in C \cap \operatorname{dom} g$ is a local minimizer for $g + \lambda h$ for some $\lambda \geq 0$. Let $x \in X$ be such that $g(x) < g(\overline{x})$. Then there is some $\varepsilon \in (0,1]$ such that for all $t \in (0,\varepsilon]$ one has $\overline{x} + t(x - \overline{x}) \in \operatorname{dom} h$ and

$$g(\overline{x}) + \lambda h(\overline{x}) \leq g\big(\overline{x} + t(x - \overline{x})\big) + \lambda h\big(\overline{x} + t(x - \overline{x})\big),$$

whence

$$\begin{aligned} 0 &\leq \frac{g\big(\overline{x} + t(x - \overline{x})\big) - g(\overline{x})}{t} + \lambda \frac{h\big(\overline{x} + t(x - \overline{x})\big) - h(\overline{x})}{t} \\ &\leq g(x) - g(\overline{x}) + \lambda \frac{h\big(\overline{x} + t(x - \overline{x})\big) - h(\overline{x})}{t}. \end{aligned}$$

Passing to the limit inferior for $t \to 0_+$, we get the contradiction $g(\overline{x}) \leq g(x)$. □

Note that the hypothesis on h is satisfied if we assume that Slater's condition (3) does not hold and h is finite-valued and Gâteaux differentiable or, even weaker, h has bilateral derivatives at any $x \in C$ and in any direction $u \in X$.

Acknowledgements. The authors thank H. H. Bauschke and J. M. Borwein for drawing their attention on the relationships among the notions and results of this paper and those contained in [4, 5, 6, 7]. They also thank one of the referees for his remarks which improved the presentation of the paper.

References

[1] V. Barbu: Nonlinear Semigroups and Differential Equations in Banach Spaces, Noordhoff, Leyden (1976).

[2] V. Barbu, Th. Precupanu: Convexity and Optimization in Banach Spaces, Editura Academiei (Bucureşti) and Sijthoff & Noordhoff Int. Publ. (Alphen aan de Rijn) (1978).

[3] V. Barbu, Th. Precupanu: Convexity and Optimization in Banach Spaces, Editura Academiei (Bucureşti) and D. Reidel Publishing Company (Dordrecht) (1986).

[4] H. H. Bauschke, J. M. Borwein: On the convergence of von Neumann's alternating projection algorithm for two sets, Set-Valued Anal. 1 (1993) 185–212.

[5] H. H. Bauschke, J. M. Borwein: On projection algorithms for solving convex feasibility problems, SIAM Rev. 38 (1996) 367–426.

[6] H. H. Bauschke, J. M. Borwein, W. Li: Strong conical hull intersection property, bounded linear regularity and Jameson's property (G), and error bounds in convex optimization, Math. Programming 86 (1999) 135–160.

[7] H. H. Bauschke, J. M. Borwein, P. Tseng: Bounded linear regularity, strong CHIP, and CHIP are distinct properties, J. Convex Analysis 7(2) (2000) 395–412.

[8] F. Bonnans, A. Shapiro: Perturbation Analysis of Optimization Problems, Springer, New York (2000).

[9] J. M. Borwein, J. D. Vanderwerff: Convex functions of Legendre type in general Banach spaces, J. Convex Analysis 8(2) (2001) 569–581.

[10] O. Cornejo, A. Jourani, C. Zălinescu: Conditioning and upper-Lipschitz inverse subdifferentials in nonsmooth optimization problems, J. Optim. Theory Appl. 95 (1997) 127–148.

[11] S. Deng: Global error bounds for convex inequality systems in Banach spaces, SIAM J. Control Optim. 36 (1998) 1240–1249.

[12] J. Diestel: Geometry of Banach Spaces, Lecture Notes in Math. 485, Springer-Verlag, Berlin (1975).

[13] I. Ekeland, R. Temam: Analyse Convexe et Problèmes Variationnels, Collection Études Mathématiques, Dunod (1974).

[14] J. R. Giles: Convex Analysis with Application in the Differentiation of Convex Functions, Research Notes in Mathematics 58, Pitman Advanced Publishing Program, Boston, Mass. (1982).

[15] J.-B. Hiriart-Urruty, Cl. Lemarechal: Convex Analysis and Minimization Algorithms, Springer Verlag, Berlin (1993).

[16] D. Klatte, W. Li: Asymptotic constraint qualifications and global error bounds for convex inequalities, Math. Programming 84 (1999) 137–160.

[17] P. J. Laurent: Approximation et Optimisation, Hermann, Paris (1972).

[18] B. Lemaire: Bounded diagonally stationary sequences in convex optimization, J. Convex Analysis 1(1) (1994) 75–86.

[19] W. Li: Abadie's constraint qualification, metric regularity and error bounds for differentiable convex inequalities, SIAM J. Control Optimiz. 7 (1998) 966–978.

[20] A. S. Lewis, J.-S. Pang: Error bounds for convex inequality systems, Proc. 5th Int. Symp. on Generalized Convexity held in Luminy, June 17–21, 1996, J.-P. Crouzeix, J.-E. Martinez-Legaz, M. Volle (eds.), Kluwer Academic Publishers (Dordrecht) (1998) 75–110.

[21] J.-S. Pang: Error bounds in mathematical programming, Math. Programming 79 (1997) 299–332.

[22] J.-P. Penot, C. Zălinescu: Harmonic sum and duality, J. Convex Analysis 7(1) (2000) 95–113.

[23] B. N. Pshenichnyi: Convex Analysis and Extremal Problems (Russian), Nauka, Moskow (1980).

[24] S. M. Robinson: An application of error bounds for convex programming in a linear space, SIAM J. Control 13 (1975) 271–273.

[25] R. T. Rockafellar: Convex Analysis, Princeton Mathematical Series 28, Princeton University Press, Princeton, New Jersey (1970)

[26] D. Tiba: On the convex programming, Preprint Nr. 2, Institute of Mathematics, Romanian Academy, Bucharest (1995).

[27] D. Tiba: Sur l'existence des multiplicateurs de Lagrange, Preprint Nr. 2, Institute of Mathematics, Romanian Academy, Bucharest (1996).

[28] D. Tiba, M. Bergounioux: General optimality conditions for constrained convex control problems, SIAM J. Control Optimiz. 34(2) (1996) 698–711.

[29] V. M. Tikhomirov: Convex analysis, in: Analysis II, R. V. Gamkrelidze (ed.), Encyclopaedia of Mathematical Sciences 14, Springer Verlag, Berlin (1990) 1–92.

[30] C. Zălinescu: Weak sharp minima, well-behaving functions and global error bounds for convex inequalities in Banach spaces, Proc. 12th Baikal Int. Conf. on Optimization Methods and their Applications, Irkutsk, Russia, 2001, Irkutsk Univ. Press 272–284.

[31] C. Zălinescu: Convex Analysis in General Vector Spaces, World Scientific, Singapore (2002).