SIAM J. OPTIM.
Vol. 7, No. 2, pp. 508–518, May 1997

013

ON UNIQUENESS OF LAGRANGE MULTIPLIERS IN OPTIMIZATION PROBLEMS SUBJECT TO CONE CONSTRAINTS*

ALEXANDER SHAPIRO†

Abstract. In this paper we study uniqueness of Lagrange multipliers in optimization problems subject to cone constraints. The main tool in our investigation of this question will be a calculus of dual (polar) cones. We give sufficient and in some cases necessary conditions for uniqueness of Lagrange multipliers in general Banach spaces. General results are then applied to two particular examples of the semidefinite and semi-infinite programming problems, respectively.

Key words. Lagrange multipliers, cone constraints, first-order optimality conditions, semidefinite programming, semi-infinite programming

AMS subject classifications. 90C30, 90C34, 90C48

PII. S1052623495279785

1. Introduction.

Consider the following optimization problem:

$$\min_{x\in X} f(x) \quad \text{subject to} \quad g(x) \in K. \tag{1.1}$$

Here X and Y are (real) Banach spaces, $f : X \to \mathbb{R}$ and $g : X \to Y$ are continuously differentiable functions, $K \subset Y$ is a convex closed cone, and

$$L(x, \lambda) = f(x) + \langle \lambda, g(x)\rangle$$

is the Lagrangian function. The first-order necessary conditions for a feasible point x_0 to be a locally optimal solution of the above problem can be written as follows (see [6, 9, 10]). Under a constraint qualification there exists $\lambda \in K^-$ such that

$$D_x L(x_0, \lambda) = 0, \tag{1.2}$$

$$\langle \lambda, g(x_0)\rangle = 0. \tag{1.3}$$

In this paper we discuss uniqueness of Lagrange multipliers satisfying the first-order necessary conditions. The question of uniqueness of Lagrange multipliers arises naturally, for example, in sensitivity analysis of optimization problems (see, e.g., [7, 13]) and in convergence analysis of Newton type optimization algorithms (cf. [2]). In case the space Y is finite dimensional and the cone K is polyhedral, there are reasonably simple necessary and sufficient conditions ensuring uniqueness of Lagrange multipliers [5]. The situation is considerably more subtle in the general case of cone constraints.

The main tool in our investigation of this question will be a calculus of dual cones. For the reader's convenience and in order to make the paper self contained we describe in the remainder of this section a few required facts from the theory of dual cones. We view the Banach space Y and its dual Y^* as paired spaces. By $\langle \alpha, y\rangle$ we denote the value $\alpha(y)$ of a continuous linear functional $\alpha \in Y^*$. We consider $\langle \cdot, \cdot\rangle$ as a bilinear form on $Y^* \times Y$ and equip Y and Y^* with a pair of compatible topologies. That is, for

* Received by the editors January 11, 1995; accepted for publication (in revised form) September 26, 1995.
http://www.siam.org/journals/siopt/7-2/27978.html
† School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0205 (ashapiro@isye.gatech.edu).

every $\alpha \in Y^*$ the linear functional $\langle \alpha, \cdot \rangle$ is continuous in the considered topology of Y, and all continuous linear functionals on Y can be represented in such form. Similarly, all linear, continuous in the considered topology of Y^* functionals can be represented in the form $\langle \cdot, v \rangle$ for some $v \in Y$. The pair of compatible topologies that we use in this paper will be the norm topology of Y and the weak star topology (w^*-topology) of Y^*.

For a cone $C \subset Y$, its polar (negative dual) cone C^- is defined as follows:

$$C^- = \{\alpha \in Y^* : \langle \alpha, y \rangle \leq 0 \text{ for all } y \in C\}.$$

Similarly, for a cone $\Sigma \subset Y^*$, its polar cone is given by

$$\Sigma^- = \{y \in Y : \langle \alpha, y \rangle \leq 0 \text{ for all } \alpha \in \Sigma\}.$$

Note that the polar cones C^- and Σ^- are always convex and closed in the considered compatible topologies; i.e., C^- is closed in the w^*-topology of Y^* and Σ^- is closed in the norm topology of Y. If C is a linear space, then C^- coincides with the orthogonal complement $C^\perp$ of C. In particular, if $\lambda \in Y^*$, then $[\lambda]^- = [\lambda]^\perp = \text{Ker}\lambda$, where $\text{Ker}\lambda$ is the null space of λ and $[\lambda]$ denotes the (one-dimensional) space generated by λ.

It follows from the Hahn–Banach theorem that if the cone $C \subset Y$ is convex, then $(C^-)^- = \text{cl}\{C\}$, where $\text{cl}\{C\}$ denotes the topological closure, in the norm topology of Y, of the cone C (e.g., [1, Chapter 1, section 5]). Similarly, if the cone $\Sigma \subset Y^*$ is convex, then $(\Sigma^-)^- = \text{cl}^*\{\Sigma\}$, where $\text{cl}^*\{\Sigma\}$ denotes the topological closure of Σ in the w^*-topology of Y^*. Note that if the space Y is reflexive and Σ is convex, then $\text{cl}^*\{\Sigma\} = \text{cl}\{\Sigma\}$.

It is straightforward to verify (cf. [1]) that if C_1 and C_2 are two cones in Y or in Y^*, then

$$(C_1 + C_2)^- = C_1^- \cap C_2^-. \tag{1.4}$$

It follows from (1.4) that the polar of the cone $C_1^- \cap C_2^-$ coincides with the polar of $(C_1 + C_2)^-$. Consequently, if the cones C_1 and C_2 are convex, then the polar cone of $C_1^- \cap C_2^-$ is given by the topological closure of the cone $C_1 + C_2$. Denote $K_1 = C_1^-$ and $K_2 = C_2^-$. It follows that $(K_1 \cap K_2)^-$ coincides with the topological closure of the cone $K_1^- + K_2^-$. Since any convex closed cone can be represented as the polar cone, we obtain that if K_1 and K_2 are two convex cones in Y or Y^*, closed in the respective compatible topology, then

$$(K_1 \cap K_2)^- = \begin{cases} \text{cl}^*\{K_1^- + K_2^-\} & \text{if } K_1, K_2 \subset Y, \\ \text{cl}\{K_1^- + K_2^-\} & \text{if } K_1, K_2 \subset Y^*. \end{cases} \tag{1.5}$$

(See, e.g., [3] for details.)

Now let S be a convex set in Y or Y^* and $v \in S$. We denote by $\mathcal{R}(S, v)$ the *radial* cone of S at v. That is, $\mathcal{R}(S, v)$ is the cone generated by the set $S - v$ or (equivalently) is the set formed by such vectors u that $v + tu \in S$ for some $t > 0$. If S is a convex cone, then $\mathcal{R}(S, v) = S + [v]$, where $[v]$ denotes the one-dimensional linear space generated by vector v. The topological closure in the norm topology of $\mathcal{R}(S, v)$ is called the *tangent* cone to S at v and denoted $T(S, v)$. When $S \subset Y^*$, we also consider $T^*(S, v) = \text{cl}^*\{\mathcal{R}(S, v)\}$. Since the radial cone of a convex set is convex, we have that if Y is reflexive, then $T^*(S, v) = T(S, v)$. If S is a convex cone closed in the respective compatible topology, we have that

$$T^*(S, v)^- = T(S, v)^- = \mathcal{R}(S, v)^- = (S + [v])^- = S^- \cap [v]^\perp.$$

Let $A : X \to Y$ be a continuous linear operator. Its adjoint operator $A^* : Y^* \to X^*$ is defined by the relation

$$\langle A^*\lambda, x\rangle = \langle \lambda Ax\rangle \text{ for all } x \in X \text{ and } \lambda \in Y^*.$$

Note that it follows from the above definition that $A^*\lambda = 0$ iff $\langle \lambda, Ax\rangle = 0$ for all $x \in X$. Therefore $\text{Ker}A^* = (AX)^\perp$.

For a convex set $S \subset Y$ we denote by $\text{int}(S)$, $\text{lin}(S)$, and $\text{ri}(S)$ its interior, the linear space generated by S, and its relative interior, respectively. That is, $\text{lin}(S)$ is the intersection of all linear subspaces which contain S and $\text{ri}(S)$ is the interior of S relative to $\text{lin}(S)$; i.e., $y \in \text{ri}(S)$ iff $y \in S$ and there is a neighborhood N (in the norm topology of Y) of y such that $N \cap \text{lin}(S) \subset S$.

PROPOSITION 1.1. *Let Y be a normed space, $C \subset Y$ be a convex cone with a nonempty interior, and L be a linear subspace of Y. Then $\text{cl}\{L + C\} = Y$ if and only if $L \cap \text{int}(C) \neq \emptyset$.*

Proof. Suppose that $L \cap \text{int}(C) \neq \emptyset$. This means that there exist $y \in L$ and a ball $B \subset Y$ of radius $r > 0$ and centered at zero such that $y + B \subset C$. We have then that $B = (-y) + y + B \subset L + C$, and since $L + C$ is a cone, it follows that $tB \subset L + C$ for any $t \geq 0$. This implies that $L + C = Y$.

Conversely, suppose that $L \cap \text{int}(C) = \emptyset$. Then by a separation theorem (e.g., [4, p. 163]) there exists $\alpha \in Y^*$, $\alpha \neq 0$ such that $\langle \alpha, y\rangle = 0$ for any $y \in L$ and $\langle \alpha, y\rangle \leq 0$ for any $y \in C$. It follows that $\alpha \in (L + C)^-$ and hence $\text{cl}\{L + C\} \subset \{y : \langle \alpha, y\rangle \leq 0\} \neq Y$. ☐

2. Basic results. Let $\lambda_0 \in K^-$ be a Lagrange multiplier satisfying optimality conditions (1.2) and (1.3). In this section we discuss general conditions for uniqueness of this Lagrange multiplier. Consider the set

$$C = \{\lambda \in K^- : \langle \lambda, g(x_0)\rangle = 0\}.$$

Note that C is a convex cone, closed in the w^*-topology of Y^*, and that $\lambda_0 \in C$. Moreover, by (1.5), $C^- = \text{cl}\{K + [g(x_0)]\}$ and hence $C^- = T(K, g(x_0))$.

PROPOSITION 2.1. *The Lagrange multiplier λ_0 is unique if and only if*

$$\mathcal{R}(C, \lambda_0) \cap [Dg(x_0)X]^\perp = \{0\}. \tag{2.1}$$

Proof. Consider a vector $\lambda \in K^-$ and let $\mu = \lambda - \lambda_0$. We have that λ satisfies (1.2) iff $[Dg(x_0)]^*\mu = 0$, and λ satisfies (1.3) iff $\lambda_0 + \mu \in C$. Therefore λ can be a Lagrange multiplier different from λ_0 iff there exists a nonzero vector $\mu \in Y^*$ such that $\mu \in [Dg(x_0)X]^\perp$ and $\mu \in \mathcal{R}(C, \lambda_0)$. ☐

In the following theorem we give sufficient, and in some cases necessary, conditions for uniqueness of λ_0 which can be viewed as dual to (2.1).

THEOREM 2.2. *The following condition is sufficient for uniqueness of λ_0:*

$$\text{cl}\{Dg(x_0)X + T(K, g(x_0)) \cap \text{Ker}\lambda_0\} = Y. \tag{2.2}$$

If $\mathcal{R}(C, \lambda_0) = T^*(C, \lambda_0)$, then condition (2.2) *is also necessary.*

Proof. Consider the cone

$$Q = Dg(x_0)X + T(K, g(x_0)) \cap \text{Ker}\lambda_0.$$

Its polar cone is given by

$$Q^- = [Dg(x_0)X]^- \cap [T(K, g(x_0)) \cap \text{Ker}\lambda_0]^-.$$

Moreover, we have that $[Dg(x_0)X]^- = [Dg(x_0)X]^\perp$ and, by (1.5),

$$[T(K, g(x_0)) \cap \mathrm{Ker}\lambda_0]^- = \mathrm{cl}^*\{[T(K, g(x_0))]^- + [\lambda_0]\} = \mathrm{cl}^*\{C + [\lambda_0]\} = T^*(C, \lambda_0).$$

Therefore,

$$Q^- = [Dg(x_0)X]^\perp \cap T^*(C, \lambda_0). \tag{2.3}$$

Suppose now that condition (2.2) holds. Then $Q^- = \{0\}$ and since $\mathcal{R}(C, \lambda_0) \subset T^*(C, \lambda_0)$, condition (2.1) follows from (2.3). Moreover, if $\mathcal{R}(C, \lambda_0) = T^*(C, \lambda_0)$, then by (2.3) condition (2.1) is equivalent to $Q^- = \{0\}$, which in turn is equivalent to (2.2). ∎

Consider now the cone $K_0 = K \cap \mathrm{Ker}\lambda_0$. We have that $T(K_0, g(x_0)) \subset T(K, g(x_0)) \cap \mathrm{Ker}\lambda_0$ and hence it follows from Theorem 2.2 that the condition

$$Dg(x_0)X + T(K_0, g(x_0)) = Y \tag{2.4}$$

is sufficient for uniqueness of λ_0. Condition (2.4) is equivalent to a constraint qualification, with respect to the cone K_0, in the sense of Robinson [11]. Its sufficiency for uniqueness of λ_0 was discussed in [12]. If the space Y is finite dimensional and the cone K is polyhedral, the condition (2.4) is also necessary (cf. [5]).

Let us remark that since $[Dg(x_0)X]^\perp = \mathrm{Ker}[Dg(x_0)]^*$, condition (2.1) is equivalent to

$$\{\mu \in \mathcal{R}(C, \lambda_0) : [Dg(x_0)]^*\mu = 0\} = \{0\}. \tag{2.5}$$

Similarly and because of (2.3), condition (2.2) is equivalent to

$$\{\mu \in T^*(C, \lambda_0) : [Dg(x_0)]^*\mu = 0\} = \{0\}. \tag{2.6}$$

In some applications it will be convenient to formulate the sufficient condition (2.2) of Theorem 2.2 in the following form.

PROPOSITION 2.3. *Let $\mathcal{L}$ be a linear space generated by the cone $\mathcal{T} = T(K, g(x_0)) \cap \mathrm{Ker}\lambda_0$ and suppose that $\mathcal{T}$ has a nonempty relative interior (relative to $\mathcal{L}$). Then condition* (2.2) *holds if the following two conditions are satisfied:*

(i) $\mathrm{cl}\{Dg(x_0)X + \mathcal{L}\} = Y$, *and*

(ii) *there exists a vector $h \in X$ such that $Dg(x_0)h \in \mathrm{ri}(\mathcal{T})$.*

Conversely, if condition (2.2) *holds and $\mathcal{L} \subset Dg(x_0)X + \mathcal{T}$, then conditions* (i) *and* (ii) *follow.*

Proof. Suppose that the above conditions (i) and (ii) are satisfied. By Proposition 1.1 it follows from condition (ii) that $\mathcal{L} \subset Dg(x_0)X + \mathcal{T}$. Together with condition (i) this implies that $\mathrm{cl}\{Dg(x_0)X + \mathcal{T}\} = Y$, meaning that condition (2.2) holds.

Conversely, let us suppose that condition (2.2) holds. Since $\mathrm{cl}\{Dg(x_0)X + \mathcal{T}\} \subset \mathrm{cl}\{Dg(x_0)X + \mathcal{L}\}$, condition (i) then follows. Also, we have that $\mathcal{L} \subset Dg(x_0)X + \mathcal{T}$ and, since $\mathcal{T} \subset \mathcal{L}$, we obtain that $\mathcal{L} = \mathcal{M} + \mathcal{T}$, where $\mathcal{M} = \mathcal{L} \cap Dg(x_0)X$. By Proposition 1.1, condition (ii) then follows. ∎

By Theorem 2.2 we obtain then that conditions (i) and (ii) of Proposition 2.3 are sufficient for uniqueness of λ_0. Note that if $Dg(x_0)X + \mathcal{T}$ is closed, then the condition $\mathcal{L} \subset Dg(x_0)X + \mathcal{T}$ follows from condition (2.2). In that case conditions (i) and (ii) are equivalent to condition (2.2).

3. Examples and applications. In this section we discuss two examples of semidefinite and semi-infinite programming. Let us start with the example of semidefinite programming. Let $X = \mathbb{R}^m$ and $Y = \mathcal{S}_n$, where $\mathcal{S}_n$ denotes the space of an $n \times n$ symmetric matrix. We equip $\mathbb{R}^m$ with the standard scalar product $x \cdot y = \sum_{i=1}^m x_i y_i$ and $\mathcal{S}_n$ with the scalar product $A \bullet B = \text{tr}AB$ for any $A, B \in \mathcal{S}_n$. The spaces X and Y can be then identified with their duals X^* and Y^*, respectively. In the space $\mathcal{S}_n$ we consider the cone K of positive semidefinite matrices, i.e., $K = \{A \in \mathcal{S}_n : A \succeq 0\}$. The cone K is convex and closed and its polar cone K^- is formed by negative semidefinite matrices, i.e., $K^- = \{\Omega \in \mathcal{S}_n : \Omega \preceq 0\}$. In what follows we denote by E^T the transpose of a matrix E.

Let $f : \mathbb{R}^m \to \mathbb{R}$ and $G : \mathbb{R}^m \to \mathcal{S}_n$ be continuously differentiable functions, $L(x, \Lambda) = f(x) + \Lambda \bullet G(x)$, and let $x_0 \in \mathbb{R}^m$ be a point satisfying the corresponding first-order optimality conditions. That is, $G(x_0) \in K$ and there exists a matrix $\Lambda_0 \in K^-$ such that

$$D_x L(x_0, \Lambda_0) = 0, \tag{3.1}$$

$$\Lambda_0[G(x_0)] = 0. \tag{3.2}$$

Note that since $G(x_0) \succeq 0$ and $\Lambda_0 \preceq 0$, condition (3.2) is equivalent to the complementarity condition $\Lambda_0 \bullet G(x_0) = 0$.

Let $r = \text{rank}G(x_0)$ and let E be an $n \times (n - r)$ matrix of full column rank $n - r$ such that $G(x_0)E = 0$. Then it is not difficult to show (cf. [15]) that the tangent cone to K at $G(x_0)$ can be written in the form

$$T(K, G(x_0)) = \{Z \in \mathcal{S}_n : E^T Z E \succeq 0\}. \tag{3.3}$$

We also have that the cone $C = \{\Lambda \in K^- : \Lambda \bullet G(x_0) = 0\}$ is given by

$$C = \{E \Theta E^T : \Theta \in \mathcal{S}_{n-r},\ \Theta \preceq 0\}. \tag{3.4}$$

We say that the *strict complementarity* condition holds if

$$\text{rank}\,\Lambda_0 + \text{rank}G(x_0) = n. \tag{3.5}$$

The Lagrange multipliers matrix Λ_0 belongs to the cone C and hence can be represented in the form $\Lambda_0 = E\Theta_0 E^T$ for some $(n-r) \times (n-r)$ symmetric, negative semidefinite matrix Θ_0. The strict complementarity condition (3.5) means that the matrix Θ_0 is nonsingular and hence is negative definite.

Under the strict complementarity condition the radial cone $\mathcal{R}(C, \Lambda_0)$ coincides with the tangent cone $T(C, \Lambda_0)$ and is given by the linear space $\{\Omega \in \mathcal{S}_n : \Omega = E\Theta E^T : \Theta \in \mathcal{S}_{n-r}\}$. Furthermore,

$$\mathcal{R}(C, \Lambda_0)^- = \mathcal{R}(C, \Lambda_0)^\perp = T(K, G(x_0)) \cap \text{Ker}\Lambda_0 = LT(K, G(x_0)),$$

where

$$LT(K, G(x_0)) = \{Z \in \mathcal{S}_n : E^T Z E = 0\}$$

is the lineality space of the cone $T(K, G(x_0))$. Therefore we obtain from Theorem 2.2 that, under the strict complementarity condition, the Lagrange multipliers matrix Λ_0 is unique iff

$$DG(x_0)\mathbb{R}^m + LT(K, G(x_0)) = \mathcal{S}_n. \tag{3.6}$$

Equation (3.6) represents a necessary and sufficient condition for a transversality relation between the mapping G and the manifold of symmetric $n \times n$ matrices of rank r (cf. [15]). It can be written in an equivalent form as follows. The adjoint $[DG(x_0)]^* : \mathcal{S}_n \to \mathbb{R}^m$ of $DG(x_0)$ is given by

$$[DG(x_0)]^*\Omega = (\Omega \bullet G_1(x_0), \dots, \Omega \bullet G_m(x_0)), \;\; \Omega \in \mathcal{S}_n,$$

where $G_i(x_0) = \partial G(x_0)/\partial x_i$ are the $n \times n$ partial derivatives matrices of $G(x)$ at $x = x_0$. Therefore, by using (2.6), we have that (3.6) is equivalent to the condition that the m-dimensional vectors $v_{ij} = (e_i^T G_1(x_0)e_j, \dots, e_i^T G_m(x_0)e_j)$, $1 \le i \le j \le n - r$, are linearly independent. Here $e_1, \dots, e_{n-r}$ are the column vectors of the matrix E.

Suppose now that $\operatorname{rank}\Theta_0 = q < n - r$. Let $\Theta_0 = V\Phi_0 V^T$ be the spectral decomposition of Θ_0; i.e., V is an $(n-r) \times q$ matrix such that $V^TV = I_q$ and Φ_0 is a $q \times q$ negative definite (diagonal) matrix. Let U be an orthogonal complement of V, i.e., U is an $(n-r) \times (n-r-q)$ matrix such that $U^TV = 0$ and $U^TU = I_{n-r-q}$, and consider the matrices $E_1 = EV$ and $E_2 = EU$ and the cone $\mathcal{T} = T(K, G(x_0)) \cap \operatorname{Ker}\Lambda_0$. We have then that

$$\mathcal{T} = \{Z \in \mathcal{S}_n : E^T ZE \succeq 0, \; E_1^T ZE_1 = 0\}.$$

Note that the column space generated by the $n \times (n-r)$ matrix $[E_1, E_2]$ is the same as the column space generated by the matrix E. Therefore we can write the cone $\mathcal{T}$ in the form

$$\mathcal{T} = \{Z \in \mathcal{S}_n : E_1^T ZE_1 = 0, \; E_1^T ZE_2 = 0, \; E_2^T ZE_2 \succeq 0\}.$$

The linear space $\mathcal{L}$, generated by the cone $\mathcal{T}$, is then given by

$$\mathcal{L} = \{Z \in \mathcal{S}_n : E_1^T ZE_1 = 0, \; E_1^T ZE_2 = 0\}$$

and the relative interior of $\mathcal{T}$ is

$$\operatorname{ri}(\mathcal{T}) = \{Z \in \mathcal{S}_n : E_1^T ZE_1 = 0, \; E_1^T ZE_2 = 0, \; E_2^T ZE_2 \succ 0\}.$$

We now can employ conditions (i) and (ii) of Proposition 2.3 in order to derive sufficient conditions for uniqueness of the Lagrange multipliers matrix Λ_0. Let $\bar{e}_1, \dots, \bar{e}_{n-r}$ be the column vectors of the matrix $[E_1, E_2]$; i.e., $\bar{e}_1, \dots, \bar{e}_q$ are the column vectors of E_1 and $\bar{e}_{q+1}, \dots, \bar{e}_{n-r}$ are the column vectors of E_2, and consider the m-dimensional vectors $\bar{v}_{ij} = (\bar{e}_i^T G_1(x_0)\bar{e}_j, \dots, \bar{e}_i^T G_m(x_0)\bar{e}_j)$, $i, j = 1, \dots, n - r$. Then, in the present situation, conditions (i) and (ii) are equivalent to the following conditions and hence, by Theorem 2.2, are sufficient for uniqueness of Λ_0.

PROPOSITION 3.1. *The following two conditions are sufficient for uniqueness of the Lagrange multipliers matrix* Λ_0.

(i′) *Vectors* $\bar{v}_{ij}$, $(i, j) \in \mathcal{I}$, *where*

$$\mathcal{I} = \{(i, j) : i, j = 1, \dots, q, \;\; i \le j\} \cup \{(i, j) : i = 1, \dots, q, \; j = q + 1, \dots, n - r\},$$

are linearly independent.

(ii′) *There exists a vector* $h \in \mathbb{R}^m$ *such that* $h \cdot \bar{v}_{ij} = 0$, $(i, j) \in \mathcal{I}$, *and*

$$\sum_{k=1}^{m} h_k E_2^T G_k(x_0) E_2 \succ 0. \tag{3.7}$$

In a sense conditions (i′) and (ii′) can be viewed as an analog of the strong Mangasarian–Fromovitz constraint qualification used in [5] for nonlinear programming problems.

Let us discuss now the example of semi-infinite programming. Consider the following optimization problem:

$$\min_{x\in\mathbb{R}^m} f(x) \quad \text{subject to} \;\; h(x,t) \le 0, \; t \in T, \tag{3.8}$$

where $f : \mathbb{R}^m \to \mathbb{R}$, $h : \mathbb{R}^m \times T \to \mathbb{R}$ and T is a compact metric space. We assume that $f(\cdot)$ and $h(\cdot, t)$ for all $t \in T$ are continuously differentiable and that $h(x,t)$ and $\nabla h(x,t)$ are continuous on $\mathbb{R}^m \times T$. (The gradient $\nabla h(x,t)$ is taken with respect to x.)

In order to formulate the inequality constraints of the semi-infinite program (3.8) in a form of cone constraints, we proceed as follows. Consider the space $C(T)$ of continuous functions $y : T \to \mathbb{R}$, equipped with the sup-norm $\|y\| = \sup_{t\in T} |y(t)|$, and the cone

$$K = \{y \in C(T) : y(t) \le 0, \; t \in T\}$$

formed by nonpositive valued continuous functions. Consider also the mapping $g : \mathbb{R}^m \to C(T)$ taking a point $x \in \mathbb{R}^m$ into the function $y = g(x)$, $y(\cdot) = h(x, \cdot)$. Then the feasible set of the program (3.8) can be defined by the cone constraint $g(x) \in K$. Note that under the above assumptions the mapping g is continuously differentiable and $[Dg(x)v](\cdot) = v \cdot \nabla h(x, \cdot)$.

The dual space Y^* of the Banach space $Y = C(T)$ is the space of finite signed measures on $(T, \mathcal{B})$, where $\mathcal{B}$ is the Borel σ-algebra of T, with the norm given by the total variation of the corresponding measure, and $\langle \lambda, y\rangle = \int_T y(t)\lambda(dt)$, $\lambda \in Y^*$, $y \in Y$. The polar cone K^- of the cone K is formed by the set of (nonnegative) Borel measures on T. For a feasible point x (satisfying $g(x) \in K$), denote by $\Delta(x)$ the set

$$\Delta(x) = \{t \in T : h(x,t) = 0\}$$

of active-at-x constraints. Then the tangent cone to K at $g(x)$ can be written in the form (e.g., [14])

$$T(K, g(x)) = \{y \in C(T) : y(t) \le 0 \;\text{ for all } t \in \Delta(x)\}. \tag{3.9}$$

Let x_0 be a locally optimal solution of (3.8). Suppose that there exists a vector $v \in \mathbb{R}^m$ such that

$$v \cdot \nabla h(x_0, t) < 0 \;\text{ for all }\; t \in \Delta(x_0). \tag{3.10}$$

In case the set T is finite, this is the Mangasarian–Fromovitz constraint qualification [8]. In the case of semi-infinite programming this condition is equivalent (e.g., [14]) to regularity of x_0 (with respect to the mapping g and the cone K) in the sense of Robinson [10].

Under the constraint qualification (3.10), x_0 corresponds with a Lagrange multiplier $\mu \in K^-$, satisfying the first-order optimality conditions, and the set of such Lagrange multipliers is bounded in the norm topology of Y^* (e.g., [9]). In the present case of semi-infinite programming, $\mu \in K^-$ is a measure and the first-order optimality conditions (1.2) and (1.3) take the form

$$\nabla f(x_0) + \int_T \nabla h(x_0, t)\mu(dt) = 0, \tag{3.11}$$

and the support of the measure μ is contained in the set $\Delta(x_0)$. Moreover, the measure μ can be chosen to be a discrete measure. That is, there are points $t_i \in \Delta(x_0)$ and numbers $\lambda_i > 0$, $i = 1, \dots, n$ such that $\mu = \sum_{i=1}^{n} \lambda_i \delta(t_i)$, where $\delta(t)$ denotes the measure of mass one at the point t. The optimality condition (3.11) then takes the form

$$\nabla f(x_0) + \sum_{i=1}^{n} \lambda_i \nabla h(x_0, t_i) = 0. \tag{3.12}$$

It is not difficult to show that if a measure μ is not discrete, then it cannot be an extreme point of the set of Lagrange multipliers measures and hence cannot be unique (e.g., [14, p. 750]). Therefore, we assume subsequently that $\mu = \sum_{i=1}^{n} \lambda_i \delta(t_i)$ is a discrete measure satisfying the first-order optimality conditions.

The cone $\mathcal{T} = T(K, g(x_0)) \cap \operatorname{Ker} \mu$ can be written here in the form

$$\mathcal{T} = \{y \in C(T) : y(t) \le 0 \text{ for all } t \in \Delta(x_0),\ y(t_i) = 0,\ i = 1, \dots, n\}. \tag{3.13}$$

The linear space $\mathcal{L}$ generated by the cone $\mathcal{T}$ is given then by

$$\mathcal{L} = \{y \in C(T) : y(t_i) = 0,\ i = 1, \dots, n\}.$$

Let us observe that it is possible that the relative interior of the cone $\mathcal{T}$ (relative to the space $\mathcal{L}$) is empty. This can happen if the points $t_1, \dots, t_n$ are not isolated points of the set $\Delta(x_0)$. Consider, for example, $T = [0, 1]$ and let $h(x_0, t) = 0$ for all $t \in [0, 1]$; i.e., $\Delta(x_0) = [0, 1]$, and let $t_1 = 1/2$, $n = 1$. Then it is not difficult to see that for any function $y(\cdot)$ in $\mathcal{T}$, one can find a function $\bar{y}(\cdot)$ in $\mathcal{L}$, arbitrarily close to $y(\cdot)$ in the sup-norm topology and such that $\bar{y}(t) > 0$ for some t sufficiently close to $1/2$.

This shows that in general we cannot apply here the sufficient conditions of Proposition 2.3. Therefore we work directly with condition (2.2) of Theorem 2.2.

PROPOSITION 3.2. *The following two conditions are necessary and sufficient for uniqueness of the Lagrange multipliers measure* $\mu = \sum_{i=1}^{n} \lambda_i \delta(t_i)$.

(i$''$) *The gradient vectors* $\nabla h(x_0, t_i)$, $i = 1, \dots, n$ *are linearly independent.*

(ii$''$) *For any neighborhood* N *of the set* $\{t_1, \dots, t_n\}$ *there exists* $v \in \mathbb{R}^m$ *such that*

$$v \cdot \nabla h(x_0, t_i) = 0, \quad i = 1, \dots, n, \tag{3.14}$$

$$v \cdot \nabla h(x_0, t) < 0, \quad t \in \Delta(x_0) \setminus N. \tag{3.15}$$

Proof. Let us first show that if the set of Lagrange multipliers measures is not a singleton, then it contains at least two different *discrete* measures. We argue as follows. Consider the set Γ of measures $\gamma \in K^-$, whose support is contained in the set $\Delta(x_0)$ and such that $\|\gamma\|^* \le 1$ and

$$c \nabla f(x_0) + \int \nabla h(x_0, t) \gamma(dt) = 0 \tag{3.16}$$

for some $c \ge 0$. Here $\|\cdot\|^*$ denotes the total variation norm on the space Y^*. For a nonnegative measure $\gamma \in K^-$, we have that $\|\gamma\|^* = \gamma(T)$. Clearly, if $\gamma \in \Gamma$ and the corresponding coefficient c in (3.16) is not zero, then $c^{-1}\gamma$ is a Lagrange multipliers measure. Conversely, if λ is a nonzero Lagrange multipliers measure, then $\lambda / \|\lambda\|^* \in \Gamma$. It is not difficult to see that Γ is convex, bounded and closed in the w^*-topology subset of Y^*, and hence is w^*-compact. By the Krein–Millman theorem it follows then that Γ

coincides with the closure (in the w^*-topology) of the convex hull of its extreme points. In order to complete the arguments it will be sufficient to show now that if a measure γ is an extreme point of Γ, then it is discrete. Consider a nondiscrete, nonzero measure $\gamma \in \Gamma$. Then $\gamma = \gamma_1 + \cdots + \gamma_{m+2}$, where γ_i, $i = 1, \ldots, m+2$ are positive measures with disjoint supports. Consider vectors $b_i = \int \nabla h(x_0, t)\gamma_i(dt)$, $i = 1, \ldots, m+2$. By dimensionality arguments there exist numbers a_i, $i = 1, \ldots, m+2$, not all of them zeros, such that $|a_i| < 1$, $\sum_{i=1}^{m+2} a_i b_i = 0$ and $\sum_{i=1}^{m+2} a_i \gamma_i(T) = 0$. Consider the measures $\gamma' = \sum_{i=1}^{m+2}(1 - a_i)\gamma_i$ and $\gamma'' = \sum_{i=1}^{m+2}(1 + a_i)\gamma_i$. Clearly $\gamma', \gamma'' \in \Gamma$ and $\gamma = (\gamma' + \gamma'')/2$. Therefore γ cannot be an extreme point of Γ.

Suppose that conditions (i$''$) and (ii$''$) hold. Because of the above arguments, in order to verify uniqueness of μ it will be sufficient to show that if $\alpha \in [Dg(x_0)\mathbb{R}^m + \mathcal{T}]^-$ and α is discrete, then $\alpha = 0$. Let $\alpha \in [Dg(x_0)\mathbb{R}^m + \mathcal{T}]^-$ be a discrete measure and let S be a *finite* subset of T containing the support of α and the set $\{t_1, \ldots, t_n\}$. We can write then $\alpha = \sum_{t \in S} \alpha(t)\delta(t)$, where $\alpha(t)$ is a nonnegative valued function on the set S. Consider a function $z \in C(T)$. Because of the condition (i$''$), there exists a vector $u \in \mathbb{R}^m$ such that $u \cdot \nabla h(x_0, t_i) = z(t_i)$, $i = 1, \ldots, n$. Choose a neighborhood N of the set $\{t_1, \ldots, t_n\}$ which does not contain other points of the set S. Then, because of the assumption (ii$''$), there exists a vector v satisfying condition (3.14) and such that $v \cdot \nabla h(x_0, t) < -c$ for all $t \in S \setminus \{t_1, \ldots, t_n\}$ and some $c > 0$. Let τ be a positive number and consider the function $a(t) = (u - \tau v) \cdot \nabla h(x_0, t)$. Note that $a \in Dg(x_0)\mathbb{R}^m$, and it follows from (3.14) that $a(t_i) = z(t_i)$, $i = 1, \ldots, n$. Moreover, we can choose τ large enough such that $a(t) \geq z(t)$ for all $t \in S \setminus \{t_1, \ldots, t_n\}$. It follows then from the representation (3.13) of the cone $\mathcal{T}$ that there exists $y \in \mathcal{T}$ such that $a(t) + y(t) = z(t)$ for all $t \in S$. Since $\int_T z(t)\alpha(dt) = \sum_{t \in S} \alpha(t)z(t)$ and $z(t)$ is an arbitrary function, it follows that $\alpha(t) = 0$ for all $t \in S$ and hence $\alpha = 0$.

Now let us show that in the present situation the condition (2.2) is necessary, as well as sufficient, for uniqueness of the Lagrange multipliers measure μ. In order to show that condition (2.2) is necessary we have to verify that $\mathcal{R}(C, \mu) = T^*(C, \mu)$. For a set $A \in \mathcal{B}$, denote by $\mathcal{Z}(A)$ the set of (nonnegative) Borel measures whose support is contained in the set A. We have that $C = \mathcal{Z}(\Delta(x_0))$ and

$$\text{(3.17)} \quad \mathcal{R}(C, \mu) = \{\alpha \in Y^* : \alpha = \alpha_1 - \alpha_2,\ \alpha_1 \in \mathcal{Z}(\Delta(x_0)),\ \alpha_2 \in \mathcal{Z}(\{t_1, \ldots, t_n\})\}.$$

Consider a signed measure $\beta \in Y^* \setminus \mathcal{R}(C, \mu)$. Let $\beta = \beta^+ - \beta^-$ be the Jordan decomposition of β; i.e., β^+ and β^- are (nonnegative) Borel measures with disjoint supports T_1 and T_2, respectively. Since $\beta \notin \mathcal{R}(C, \mu)$, we have that $T_2 \not\subset \{t_1, \ldots, t_n\}$. Consequently there is a nonzero function $y \in K$ whose support has empty intersection with the set $\{t_1, \ldots, t_n\}$ and such that $\int_T y(t)\beta(dt) < 0$. It follows from the representation of $\mathcal{R}(C, \mu)$ given in (3.17) that for any $\alpha \in \mathcal{R}(C, \mu)$, $\int_T y(t)\alpha(dt) \geq 0$ and hence we can separate β from $\mathcal{R}(C, \mu)$ by the linear functional $\langle \cdot, y\rangle$. This shows that $\mathcal{R}(C, \mu)$ is closed in the w^*-topology of Y^* and hence $\mathcal{R}(C, \mu) = T^*(C, \mu)$.

Suppose now that condition (2.2) holds. Since $\mathcal{T} \subset \mathcal{L}$, condition (2.2) implies that $Dg(x_0)\mathbb{R}^m + \mathcal{L}$ is dense in $C(T)$. Therefore $\mathcal{L}^\perp \cap \operatorname{Ker}[Dg(x_0)]^* = \{0\}$ and hence condition (i$''$) follows. Furthermore, consider a function $z \in C(T)$ such that $z(t_i) = 0$, $i = 1, \ldots, n$, and $z(t) > 0$ for all $t \in T \setminus \{t_1, \ldots, t_n\}$. Let N be an open neighborhood of the set $\{t_1, \ldots, t_n\}$. Then the set $\Delta(x_0) \setminus N$ is compact and hence there exists $\varepsilon > 0$ such that $z(t) \geq \varepsilon$ for all $t \in \Delta(x_0) \setminus N$. It follows then from condition (2.2) that there exists a function $a(t) = w \cdot \nabla h(x_0, t)$ such that $a(t) \geq \varepsilon/2$ for all $t \in \Delta(x_0) \setminus N$ and $a(t_i)$, $i = 1, \ldots, n$, are arbitrarily close to zero. Because of the condition (i$''$), we can find a vector $u \in \mathbb{R}^m$ such that $u \cdot \nabla h(x_0, t_i) = a(t_i)$, $i = 1, \ldots, n$. We obtain then that $(w - u) \cdot \nabla h(x_0, t_i) = 0$, $i = 1, \dots, n$. Moreover, for $a(t_i)$, $i = 1, \dots, n$ sufficiently close to zero, we can choose such u that $(w-u) \cdot \nabla h(x_0, t) \geq \varepsilon/3$ for all $t \in \Delta(x_0) \setminus N$. Vector $v = u - w$ then satisfies (3.14) and (3.15) and hence condition (ii″) follows. □

Note that the Mangasarian–Fromovitz constraint qualification (3.10) is not assumed in Proposition 3.2. We only assume existence of a discrete Lagrange multipliers measure μ.

Vector v in the condition (ii″) of Proposition 3.2 generally depends on the neighborhood N. It is natural then to ask whether condition (ii″) can be replaced by the following stronger condition.

(ii‴) There exists $v \in \mathbb{R}^m$ such that

$$v \cdot \nabla h(x_0, t_i) = 0, \quad i = 1, \dots, n, \tag{3.18}$$

$$v \cdot \nabla h(x_0, t) < 0, \quad t \in \Delta(x_0) \setminus \{t_1, \dots, t_n\}. \tag{3.19}$$

It is not difficult to see that if the set of active constraints $\Delta(x_0)$ is *finite*, then conditions (ii″) and (ii‴) are equivalent. As the following example shows, however, in general condition (ii‴) is not necessary for uniqueness of the Lagrange multipliers measure μ.

Example 3.1. Let $T = [0, 4]$ and consider $h : \mathbb{R}^3 \times [0, 4] \to \mathbb{R}$ of the form $h(x, t) = x_1 a_1(t) + x_2 a_2(t) + x_3 a_3(t)$, with the functions $a_i(t)$ defined as follows:

$$a_1(t) = \begin{cases} t^2, & t \in [0, 1], \\ 1.5 - 0.5t, & t \in [1, 3], \\ 0, & t \in [3, 4], \end{cases}$$

$$a_2(t) = \begin{cases} -t, & t \in [0, 1], \\ t - 2, & t \in [1, 4], \end{cases}$$

and $a_3(t) = 1$ for $t \in [0, 4]$. Also let $f(x)$ be a linear function with $\nabla f(x) = (0, 0, -1)$. We have then that at $x_0 = 0$, $\nabla f(x_0) + \nabla h(x_0, 0) = 0$ and $\Delta(x_0) = [0, 4]$. Therefore the first-order optimality conditions hold at $x_0 = 0$, with the Lagrange multipliers measure $\mu = \delta(t_1)$, $t_1 = 0$, and hence, since the considered program is convex, $x_0 = 0$ is the optimal solution of the considered program. We also have that for $v = (0, 0, -1)$ and all $t \in [0, 4]$, $v \cdot \nabla h(x_0, t) = -1$ and hence condition (3.10) is satisfied.

Let us observe now that condition (ii‴) does not hold here. Indeed, suppose there is a vector $v = (v_1, v_2, v_3)$ satisfying (3.18) and (3.19). It follows then from (3.18) that $v_3 = 0$ and from (3.19) that $v_2 < 0$. We obtain that $v \cdot \nabla h(x_0, 0) = 0$ and $\partial[v \cdot \nabla h(x_0, 0)]/\partial t > 0$. Therefore $v \cdot \nabla h(x_0, t)$ is positive for sufficiently small $t > 0$, which of course contradicts (3.19).

On the other hand, it is not difficult to verify that conditions (i″) and (ii″) of Proposition 3.2 are satisfied here and hence μ is unique. This demonstrates that conditions (ii″) and (ii‴) are not equivalent and condition (ii‴) is not necessary for uniqueness of μ.

REFERENCES

[1] J. P. Aubin and I. Ekeland, *Applied Nonlinear Analysis*, Wiley, New York, 1984.
[2] J. F. Bonnans, *Local analysis of Newton-type methods for variational inequalities and nonlinear programming*, Appl. Math. Optim., 29 (1994), pp. 161–186.

[3] R. B. Holmes, *Geometric Functional Analysis and Its Applications*, Springer-Verlag, Berlin, New York, 1975.

[4] A. D. Ioffe and V. M. Tihomirov, *Theory of Extremal Problems*, North-Holland, Amsterdam, 1979.

[5] J. Kyparisis, *On uniqueness of Kuhn-Tucker multipliers in non-linear programming*, Math. Programming, 32 (1985), pp. 242–246.

[6] S. Kurcyusz, *On the existence and nonexistence of Lagrange multipliers in Banach spaces*, J. Optim. Theory Appl., 20 (1976), pp. 81–110.

[7] F. Lempio and H. Maurer, *Differential stability in infinite-dimensional nonlinear programming*, Appl. Math. Optim., 6 (1980), pp. 139–152.

[8] O. L. Mangasarian and S. Fromovitz, *The Fritz John necessary optimality conditions in the presence of equality and inequality constraints*, J. Math. Anal. Appl., 7 (1967), pp. 37–47.

[9] H. Maurer and J. Zowe, *First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems*, Math. Programming, 16 (1979), pp. 98–110.

[10] S. M. Robinson, *First order conditions for general nonlinear optimization*, SIAM J. Appl. Math., 30 (1976), pp. 597–607.

[11] S. M. Robinson, *Stability theory for systems of inequalities, Part* II: *Differentiable nonlinear systems*, SIAM J. Numer. Anal., 13 (1976), pp. 497–513.

[12] A. Shapiro, *Perturbation analysis of optimization problems in Banach spaces*, Numer. Funct. Anal. Optim., 13 (1992), pp. 97–116.

[13] A. Shapiro, *Sensitivity analysis of parametrized programs via generalized equations*, SIAM J. Control Optim., 32 (1994), pp. 553–571.

[14] A. Shapiro, *On Lipschitzian stability of optimal solutions of parametrized semi-infinite programs*, Math. Oper. Res., 19 (1994), pp. 743–752.

[15] A. Shapiro, *First and second order analysis of nonlinear semidefinite programs*, Math. Programming Series B, to appear.