Mathematical Programming 57 (1992) 1-14
North-Holland

Minimum principle sufficiency

M.C. Ferris and O.L. Mangasarian
Computer Sciences Department, University of Wisconsin, Madison, WI 53706, USA

Received 25 September 1989
Revised manuscript received 23 April 1990

We characterize the property of obtaining a solution to a convex program by minimizing over the feasible region a linearization of the objective function at any of its solution points (Minimum Principle Sufficiency). For the case of a monotone linear complementarity problem this MPS property is completely equivalent to the existence of a nondegenerate solution to the problem. For the case of a convex quadratic program, the MPS property is equivalent to the span of the Hessian of the objective function being contained in the normal cone to the feasible region at any solution point, plus the cone generated by the gradient of the objective function at any solution point. This in turn is equivalent to the quadratic program having a weak sharp minimum. An important application of the MPS property is that minimizing on the feasible region a linearization of the objective function at a point in a neighborhood of a solution point gives an exact solution of the convex program. This leads to finite termination of convergent algorithms that periodically minimize such a linearization.

Key words: Minimum principle, convex programs, linear complementarity.

1. Introduction

It is well known by the minimum principle (see [9, Theorem 9.3.3] and [17, Theorem 7.1.1]) that each solution of a convex program with a differentiable objective function minimizes the linearization of the objective at any solution point on the feasible region. Our concern in this paper is the converse: when do *all* minimizers (over the feasible region) of a linearization of the objective function at any solution point give an exact solution of the convex program? To see that this converse does not hold in general, even when the objective function is strongly convex, consider the trivial example: $\min_{x\geqslant 0} x^2$ with the unique solution $x=0$. For this problem $\operatorname{argmin}_{x\geqslant 0}\langle(\nabla f(\bar{x}), x\rangle$ is the nonnegative real line, which is the entire feasible region of the problem, rather than the optimal solution set. Curiously, however, for a solvable monotone linear complementarity problem, the converse implication holding is completely equivalent to the existence of *some* nondegenerate solution to the complementarity problem (see Theorem 13 below). We shall refer to this converse condition as the minimum principle sufficiency (MPS) property. Part of the

This material is based on research supported by National Science Foundation Grants DCR-8521228 and CCR-8723091, and Air Force Office of Scientific Research Grants AFOSR 86-0172 and AFOSR and AFOSR 89-0410.

importance of the MPS property stems from the consequence that minimizing a linearization of the objective at any point in a sufficiently small neighborhood of any solution point gives an exact solution of the convex program (see Theorem 10 and Corollary 14 below). This leads to a *finite termination* of any computational algorithm which periodically solves the linearized problem $\min_{x \in S} \langle \nabla f(x^k), x \rangle$ where $\{x^k\}$ are the iterates of the algorithm. Related results on finite termination have recently been given in [5, 4, 1, 3]. Another useful relation is the equivalence of the MPS property to the existence of a weak sharp minimum [7] for convex quadratic programs. The weak sharp minimum property, which is naturally possessed by all solvable linear programs [15], extends the finite termination property of the proximal point algorithm for linear programs [18, 2] to convex programs [7, 6].

The paper is organized as follows. In Section 2 we give various results for the MPS property for general convex and quadratic programs, and in Section 3 we specialize and sharpen these results to monotone linear complementarity problems. The monotone linear complementarity problem seems to be a particularly suitable problem for invoking the MPS property in the sense that it endows the problem with a useful quasi-linearity property which was already pointed out in [13]. In [13], the MPS property was obtained as a consequence of nondegeneracy. In this paper, we show that these two properties are equivalent to each other as well as the existence of a weak sharp minimum for the equivalent quadratic program.

The first principal result of Section 2, Theorem 3, establishes the MPS property for a general convex program with a twice differentiable objective function under the assumption that the span of the Hessian of the objective function is contained in the algebraic sum of the normal cone to the feasible region at any optimal point, plus any nonnegative multiple of the gradient of the objective function at any optimal point. This sufficient condition for the MPS property turns out to be also necessary for a convex quadratic program (Theorem 6). It is also equivalent to the existence of a weak sharp minimum for a convex quadratic program (Theorem 6). In Theorem 9 we give a simple proof of a strong-upper semicontinuity result for perturbed linear programs due to Polyak and Tretiyakov [18], which shows that if perturbations of a cost vector of a linear program converge to an unperturbed cost vector such that each perturbed problem is solvable, then for all sufficiently small but finite values of the perturbation, all solutions of the perturbed problems solve the unperturbed problem. Robinson [19] established a stronger version of this result. This result is used in Theorem 10 to show how finite termination can be achieved under the MPS property by periodically solving a problem with a linearized objective function.

In Section 3 we specialize the MPS property to linear complementarity problems (LCP's). We first show (Lemma 11) that for a feasible LCP, its linearization at any feasible point is solvable. In Theorem 12 we show that for an LCP with a nondegenerate vertex solution, the linearized problem at *any* point in a neighborhood of the vertex solution is uniquely solved by the nondegenerate vertex. The principal result of Section 3, Theorem 13, establishes the equivalence of nondegeneracy, the MPS property, the existence of a weak sharp minimum, as well a normal cone inclusion property for a monotone linear complementarity problem. Corollary 14 shows that for nondegenerate monotone linear complementarity problems, solving the linearized problem at a point in the neighborhood of *any* solution point (degenerate or not) will yield a solution to the LCP.

A brief word about notation is provided here for the reader's convenience. For a vector x in the n-dimensional real space $\mathbb{R}^n$, $\|x\|$ will denote the Euclidean norm and $\langle x, y\rangle$ will denote the scalar product of x and y in $\mathbb{R}^n$. For an $m \times n$ real matrix A signified by $A \in \mathbb{R}^{m\times n}$, A_i denotes the ith row, while A^{T} will denote the transpose. For $M \in \mathbb{R}^{n\times n}$, $\|M\|$ will denote the Euclidean norm. The identity matrix will be denoted by I while a vector of ones will be denoted by e. The closed ball of radius δ around $\bar{x}$ will be denoted by $\mathbb{B}_\delta(\bar{x})$. The normal cone $N(\bar{x}|S)$ to a convex set $S \subseteq \mathbb{R}^n$ at $\bar{x} \in S$ is defined by $\{y \mid \langle y, x-\bar{x}\rangle \leqslant 0\ \forall x \in S\}$. For a differentiable function $f: \mathbb{R}^n \to \mathbb{R}$, $\nabla f(x)$ denotes the gradient at x. For a convex set $S \subseteq \mathbb{R}^n$ the cone generated by S is defined by cone $S = \{\lambda x \mid \lambda \geqslant 0, x \in S\}$. For a matrix $B \in \mathbb{R}^{m\times n}$ we define the linear spaces span (B) and ker B by $\{z \mid z = Bu,\ u \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$ and $\{u \mid Bu = 0,\ u \in \mathbb{R}^n\}$, respectively, and the conjugate cone conj B by $\{u \mid Bu \geqslant 0,\ u \in \mathbb{R}^n\}$. If $S \subseteq \mathbb{R}^n$ and $f: \mathbb{R}^n \to \mathbb{R}$, the set $\operatorname{argmin}_{x\in S} f(x)$, denotes the (possibly empty) solution set of $\min_{x\in S} f(x)$, and if S is convex, the set arg vertex S denotes the (possibly empty) set of extreme points of S.

2. Convex and quadratic programs

We shall be concerned with the convex program

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad x \in S \end{aligned} \tag{1}$$

where S is a closed convex subset of $\mathbb{R}^n$ and $f: \mathbb{R}^n \to \mathbb{R}$ is a differentiable convex function on $\mathbb{R}^n$. We assume that the solution set of (1),

$$\bar{S} := \operatorname{argmin} f(x),$$

is nonempty. We begin by stating the following key result that will be used throughout the paper.

Theorem 1 [14]. *Let f be differentiable and convex on $\mathbb{R}^n$, let S be a closed convex subset of $\mathbb{R}^n$ and let $\bar{x} \in \bar{S}$. Then*

$$\begin{aligned} \bar{S} &= \{x \in S \mid \langle \nabla f(\bar{x}), x-\bar{x}\rangle = 0, \nabla f(x) = \nabla f(\bar{x})\} \\ &= \{x \in S \mid \langle \nabla f(\bar{x}), x-\bar{x}\rangle \leqslant 0, \nabla f(x) = \nabla f(\bar{x})\}. \quad \square \end{aligned} \tag{2}$$

As a consequence, the minimum principle (see [9, Theorem 9.3.3; 17, Theorem 17.1.1]) can be written in the following form which will be useful for our purposes.

Theorem 2. *Let f be differentiable and convex on $\mathbb{R}^n$, let S be a closed convex subset of $\mathbb{R}^n$. Then*

$$\langle \nabla f(x), z-y\rangle \geqslant 0, \quad \forall x, y \in \bar{S}, \quad \forall x \in S. \quad \square$$

It is obvious from the minimum principle that $\bar{S} \subseteq \operatorname{argmin}_{z \in S}\langle \nabla f(x), z\rangle$ for all $x \in \bar{S}$. We begin by giving a sufficient condition for the opposite inclusion for the general convex program (1). Later, we will show that this condition is also necessary for a convex quadratic program and the monotone linear complementarity problem.

Theorem 3. *Let f be a twice differentiable convex function on $\mathbb{R}^n$, let S be a closed convex subset of $\mathbb{R}^n$, let $\bar{x} \in \bar{S}$ and let*

$$H(x) := \int_{t=0}^{t=1} f(\bar{x} + t(x - \bar{x}))\, \mathrm{d}t.$$

Then

$$\operatorname{span}(H(S)) \subseteq N(\bar{x} \mid S) + \operatorname{cone} \nabla f(\bar{x}) \;\Rightarrow\; \operatorname*{argmin}_{x \in S}\langle \nabla f(\bar{x}), x\rangle \subseteq \bar{S} \tag{3}$$

where $N(\bar{x} \mid S)$ is the normal cone to S at $\bar{x}$, cone $\nabla f(\bar{x})$ *is the cone generated by $\nabla f(\bar{x})$ and* $\operatorname{span}(H(S)) := \bigcup_{x \in S} \operatorname{span}(H(x))$.

Proof.

$$\operatorname{span}(H(S)) \subseteq N(\bar{x} \mid S) + \operatorname{cone} \nabla f(\bar{x})$$

$\Leftrightarrow$ $\forall x \in S,\ \forall h \in \mathbb{R}^n,\ \exists \xi(x, h) \geq 0\colon\ \langle H(x)h - \xi \nabla f(\bar{x}), y - \bar{x}\rangle \leq 0\ \ \forall y \in S$

$\Rightarrow$ $\forall x \in S,\ \forall h \in \mathbb{R}^n,\ \exists \xi(x, h) \geq 0\colon\ \langle H(x)h - \xi \nabla f(\bar{x}), x - \bar{x}\rangle \leq 0$
(set $y = x$ in previous statement)

$\Rightarrow$ $\forall h \in \mathbb{R}^n\colon\quad \langle \nabla f(\bar{x}), x - \bar{x}\rangle \leq 0,\ \langle H(x)h, x - \bar{x}\rangle > 0,$

has no solution $x \in S$

(for if it did have a solution $x \in S$ then we would contradict the previous statement)

$\Leftrightarrow$ $\langle \nabla f(\bar{x}), x - \bar{x}\rangle \leq 0,\ \ H(x)(x - \bar{x}) \neq 0,$ has no solution $x \in S$

$\Leftrightarrow$ $\langle \nabla f(\bar{x}), x - \bar{x}\rangle \leq 0,\ \ x \in S, \nabla f(x) \neq \nabla f(\bar{x}),$ has no solution

(since $\nabla f(x) - \nabla f(\bar{x}) = H(x)(x - \bar{x})$)

$\Leftrightarrow$ $\operatorname*{argmin}_{x \in S}\langle \nabla f(\bar{x}), x\rangle \subseteq \bar{S},$

the last equivalence following from Theorem 1. □

It is interesting to note that the first inclusion of (3),

$$\operatorname{span}(H(S)) \subseteq N(\bar{x} \mid S) + \operatorname{cone} \nabla f(\bar{x}),$$

which is an extension of the minimum principle (Theorem 2),

$$0 \in N(\bar{x} \mid S) + \nabla f(\bar{x}),$$

may be interpreted as the inclusion in $N(\bar{x} \mid S) + \text{cone } \nabla f(\bar{x})$ of all possible contributions of the gradient of the "quadratic part" of f. When $f(x) = \frac{1}{2}\langle x, Hx\rangle + \langle d, x\rangle$, as in Theorem 6 below, this is merely Hx for all x.

We now show for a convex quadratic program that the backward implication of (3) also holds, and in fact the MPS property is equivalent to the existence of a weak sharp minimum [7] for the convex quadratic program. First of all, we define this notion.

Definition 4 [7]. Let $f: \mathbb{R}^n \to \mathbb{R}$ and $S \subseteq \mathbb{R}^n$ be convex and let $\bar{S} := \operatorname{argmin}_{x \in S} f(x)$ be nonempty and closed. The problem $\min_{x \in S} f(x)$ is said to have a *weak sharp minimum* if there exists a positive constant α such that

$$f(x) - f(P(x \mid \bar{S})) \geq \alpha \|x - P(x \mid \bar{S})\| \quad \forall x \in S$$

where $P(x \mid \bar{S}) \in \operatorname{argmin}_{z \in S} \|z - x\|$ and $\|\cdot\|$ is some norm on $\mathbb{R}^n$, that is the orthogonal projection of x on $\bar{S}$.

Note that all solvable linear programs have weak sharp minima [15]. Note also that a weak sharp minimum generalizes the concept of a sharp minimum [17, p. 205], in which $P(x \mid \bar{S})$ is replaced by the fixed unique solution $\bar{x}$ of $\min_{x \in S} f(x)$. The sharp minimum condition is more stringent than the weak sharp minimum condition, and does not hold, in general, even for linear programs. We also relate the MPS property to the notion of nondegeneracy for the convex problem and its dual. We shall use the following equivalent forms of the definition, which we state as an easily established lemma.

Lemma 5. *Consider the convex quadratic program* $\min_{x \in S} f(x)$ *with nonempty solution set* $\bar{S}$ *and*

$$f(x) := \tfrac{1}{2}\langle x, Hx\rangle + \langle d, x\rangle, \qquad S := \{x \mid Ax \geq b,\ x \geq 0\},$$

where $H \in \mathbb{R}^{n \times n}$ *is symmetric and positive semidefinite,* $A \in \mathbb{R}^{m \times n}$, $d \in \mathbb{R}^n$ *and* $b \in \mathbb{R}^m$. *The following are equivalent:*

(i) $\min_{x \in S} f(x)$ *and its dual have a nondegenerate primal-dual solution* $(\hat{x}, \hat{u})$ (*see* (7)).

(ii) $-\nabla f(\hat{x}) \in \operatorname{ri} N(\hat{x} \mid S)$ (*see Dunn* [5]).

(iii) *There exist* $(u, v) > 0$ *with* $\nabla f(\hat{x}) = A_{\hat{x}}^{\mathrm{T}} u + I_{\hat{x}}^{\mathrm{T}} v$ *where* $A_{\hat{x}} = \{A_i \mid A_i\hat{x} = b_i\}$ *and* $I_{\hat{x}} = \{I_i \mid \hat{x}_i = 0\}$. □

We now characterize the MPS property for convex quadratic programs.

Theorem 6. *Consider the convex quadratic program* $\min_{x \in S} f(x)$ *with nonempty solution set* $\bar{S}$ *and*

$$f(x) := \tfrac{1}{2}\langle x, Hx\rangle + \langle d, x\rangle, \qquad S = \{x \mid Ax \geq b,\ x \geq 0\}, \quad \bar{x} \in \bar{S},$$

where $H \in \mathbb{R}^{n \times n}$ *is symmetric and positive semidefinite,* $A \in \mathbb{R}^{m \times n}$, $d \in \mathbb{R}^n$ *and* $b \in \mathbb{R}^m$, *and let* $A_{\bar{x}} := \{A_i \mid A_i\bar{x} = b_i\}$ *and* $I_{\bar{x}} := \{I_i \mid \bar{x}_i = 0\}$. *Then* (i) *to* (iv) *below are equivalent* *and they imply* (v).

(i) $\operatorname{argmin}_{x\in S}\langle \nabla f(\bar{x}), x\rangle \subseteq \bar{S}$ (*MPS*).

(ii) $\min_{x\in S} f(x)$ *has a weak sharp minimum* (*WSM*).

(iii) $\operatorname{span}(H) \subseteq N(\bar{x}\,|\,S) + \operatorname{cone}\nabla f(\bar{x})$ *for each* $\bar{x}\in\bar{S}$, *or equivalently*

$$\operatorname{conj}\begin{bmatrix} A_{\bar{x}} \\ I_{\bar{x}} \\ -\nabla f(\bar{x}) \end{bmatrix} \subseteq \ker H.$$

(iv) $\operatorname{span}(H) \subseteq N(\bar{x}\,|\,S) + \operatorname{span}(\nabla f(\bar{x}))$ *for each* $\bar{x}\in\bar{S}$.

(v) $\min_{x\in S} f(x)$ *and its dual have a nondegenerate primal–dual solution* $(\hat{x}, \hat{u})$.

Furthermore (v) *implies* (i) *to* (iv) *under the assumption*

$$\operatorname{span}(H) \subseteq \operatorname{span}(N(\hat{x}\,|\,S)) + \operatorname{span}(\nabla f(\hat{x}))$$

or equivalently

$$\ker\begin{bmatrix} A_{\hat{x}} \\ I_{\hat{x}} \\ -\nabla f(\hat{x}) \end{bmatrix} \subseteq \ker H. \tag{4}$$

Proof. Note that the equivalence given in (iii) follows immediately from the Farkas Theorem [9, Theorem 2.4.6], since, for each $\bar{x}\in\bar{S}$,

$$\operatorname{span}(H) \subseteq N(\bar{x}\,|\,S) + \operatorname{cone}\nabla f(\bar{x})$$

$$\Leftrightarrow\ Hh = -A_{\bar{x}}^{\mathrm{T}}u - I_{\bar{x}}^{\mathrm{T}}v + \nabla f(\bar{x})\eta \text{ has solution } (\eta, u, v) \geq 0,\ \forall h\in\mathbb{R}^n$$

$$\Leftrightarrow\ \begin{bmatrix} A_{\bar{x}} \\ I_{\bar{x}} \\ -\nabla f(\bar{x}) \end{bmatrix} x \geq 0,\quad \langle h, Hx\rangle > 0, \text{ has no solution } x,\ \forall h\in\mathbb{R}^n$$

(by Farkas Theorem [9, Theorem 2.4.6])

$$\Leftrightarrow\ \begin{bmatrix} A_{\bar{x}} \\ I_{\bar{x}} \\ -\nabla f(\bar{x}) \end{bmatrix} x \geq 0 \ \Rightarrow\ Hx = 0$$

$$\Leftrightarrow\ \operatorname{conj}\begin{bmatrix} A_{\bar{x}} \\ I_{\bar{x}} \\ -\nabla f(\bar{x}) \end{bmatrix} \subseteq \ker H.$$

(i)$\Leftrightarrow$(iii) By Theorem 1, it follows that $\nabla f(\bar{x})$ is constant on the solution set of a convex program, and hence for any $\bar{x}\in\bar{S}$,

$$\operatorname*{argmin}_{x\in S}\langle \nabla f(\bar{x}), x\rangle \subseteq \bar{S}$$

$$\Leftrightarrow\ \langle \nabla f(\bar{x}), x-\bar{x}\rangle \leq 0,\quad Ax \geq b,\ x\geq 0,\quad \nabla f(x) - \nabla f(\bar{x}) = H(x-\bar{x}) \neq 0,$$
has no solution x

(by Theorems 1 and 2)

$$\Leftrightarrow \langle \nabla f(\bar{x}), x-\bar{x}\rangle \leq 0, \quad Ax \geq b, x \geq 0, \quad \langle h, H(x-\bar{x})\rangle > 0,$$
has no solution x, $\forall h \in \mathbb{R}^n$

$$\Leftrightarrow -\nabla f(\bar{x}) + A^{\mathrm{T}}u + v + \xi Hh = 0, \quad \langle \nabla f(\bar{x}), \bar{x}\rangle\eta - \langle b, u\rangle - \langle h, H\bar{x}\rangle\xi + \rho = 0,$$
has solution $(\eta, u, v) \geq 0, 0 \neq (\xi, \rho) \geq 0$, $\forall h \in \mathbb{R}^n$

(by Motzkin's Theorem [9, Theorem 2.4.2])

$$\Leftrightarrow -\nabla f(\bar{x})\eta + A^{\mathrm{T}}u + v + Hh = 0, \quad \langle \nabla f(\bar{x}), \bar{x}\rangle\eta - \langle b, u\rangle - \langle h, H\bar{x}\rangle \leq 0,$$
has solution $(\eta, u, v) \geq 0$, $\forall h \in \mathbb{R}^n$

(Set $\xi = 1$. For, if $\xi = 0$, setting $\eta = 0$ contradicts primal feasibility, while setting $\eta > 0$ contradicts the fact that $\langle \nabla f(\bar{x}), \bar{x}\rangle = \min_{x\in S}\langle \nabla f(\bar{x}), x\rangle$.)

$$\Leftrightarrow -\nabla f(\bar{x})\eta + A^{\mathrm{T}}u + v + Hh = 0, \quad 0 \leq \langle u, A\bar{x} - b\rangle + \langle v, \bar{x}\rangle \leq 0,$$
has solution $(\eta, u, v) \geq 0$, $\forall h \in \mathbb{R}^n$

(substitute for Hh from the equality in the inequality)

$$\Leftrightarrow Hh = -A_{\bar{x}}^{\mathrm{T}}u - I_{\bar{x}}^{\mathrm{T}}v + \nabla f(\bar{x})\eta$$
has solution $(\eta, u, v) \geq 0$, $\forall h \in \mathbb{R}^n$

$$\Leftrightarrow \operatorname{span}(H) \subseteq N(\bar{x} \mid S) + \operatorname{cone} \nabla f(\bar{x}) \quad \text{for each } \bar{x} \in \bar{S}.$$

(iii)$\Leftrightarrow$(iv) The forward implication is trivial. For the backward implication we have by Theorem 2 that

$$0 \in N(\bar{x} \mid S) + \nabla f(\bar{x}).$$

Combining this with

$$\operatorname{span}(H) \subseteq N(\bar{x} \mid S) + \operatorname{span}(\nabla f(\bar{x}))$$

gives (iii).

(ii)$\Leftrightarrow$(iv) See [3].

(i)$\Rightarrow$(v) We need to show that the dual quadratic programs [9, Problem 8.3.9]

$$\min_x\{\tfrac{1}{2}\langle x, Hx\rangle + \langle d, x\rangle \mid Ax \geq b, x \geq 0\} \tag{5}$$

and

$$\max_x\{-\tfrac{1}{2}\langle x, Hx\rangle + \langle b, u\rangle \mid Hx - A^{\mathrm{T}}u + d \geq b, u \geq 0\} \tag{6}$$

have a nondegenerate primal-dual solution $(\hat{x}, \hat{u})$, that is

$$\hat{x} + H\hat{x} - A^{\mathrm{T}}\hat{u} + d > 0, \qquad \hat{u} + A\hat{x} - b > 0. \tag{7}$$

Let $\tilde{x} \in \bar{S}$. By the nondegeneracy result for linear programming [8, Corollary 2A], the dual linear programs

$$\min_x \{\langle \nabla f(\tilde{x}), x\rangle \mid Ax \geq b, x \geq 0\}$$

and

$$\max_u \{\langle b, u\rangle \mid A^{\mathrm{T}}u \leq \nabla f(\tilde{x}), u \geq 0\}$$

have a nondegenerate primal-dual solution $(\hat{x}, \hat{u})$, that is

$$\begin{aligned} &A\hat{x} - b \geq 0, \quad \hat{x} \geq 0, \quad \langle \hat{u}, A\hat{x} - b\rangle = 0, \quad \hat{u} + A\hat{x} - b > 0, \\ &-A^{\mathrm{T}}\hat{u} + \nabla f(\tilde{x}) \geq 0, \ \hat{u} \geq 0, \ \langle \hat{x}, -A^{\mathrm{T}}\hat{u} + \nabla f(\tilde{x})\rangle = 0, \ \hat{x} - A^{\mathrm{T}}\hat{u} + \nabla f(\tilde{x}) > 0. \end{aligned} \tag{8}$$

By hypothesis of (i), $\hat{x} \in \bar{S}$, and since $\tilde{x} \in \bar{S}$, it follows by Theorem 1 that $\nabla f(\tilde{x}) = \nabla f(\hat{x})$, and therefore $\nabla f(\tilde{x})$ can be replaced by $\nabla f(\hat{x})$ in (8). With this replacement conditions (8) are sufficient Karush–Kuhn–Tucker condtions for $\hat{x}$ to solve the primal problem (5) and $(\hat{x}, \hat{u})$ satisfy the nondegeneracy conditions. Since $(\hat{x}, \hat{u})$ is feasible for (6) and

$$\begin{aligned} (\tfrac{1}{2}\langle \hat{x}, H\hat{x}\rangle + \langle d, \hat{x}\rangle) - (-\tfrac{1}{2}\langle \hat{x}, H\hat{x}\rangle + \langle b, \hat{u}\rangle) &= \langle \hat{x}, H\hat{x}\rangle + \langle d, \hat{x}\rangle) - \langle b, \hat{u}\rangle \\ &= \langle \hat{x}, H\hat{x} - A^{\mathrm{T}}\hat{u} + d\rangle + \langle \hat{u}, A\hat{x} - b\rangle \\ &= 0, \end{aligned}$$

it follows that $(\hat{x}, \hat{u})$ is also optimal for the dual (6).

(v)$\Rightarrow$(iii) We establish this implication assuming (4) holds.

$$\begin{aligned} \text{(v)} \ &\Leftrightarrow \ \nabla f(\hat{x}) - A_{\hat{x}}^{\mathrm{T}}u - I_{\hat{x}}^{\mathrm{T}}v = 0 \ \text{ has solution } (u, v) > 0 \\ &\qquad \text{(by Lemma 5)} \\ &\Leftrightarrow \ \nabla f(\hat{x})\xi - A_{\hat{x}}^{\mathrm{T}}u - I_{\hat{x}}^{\mathrm{T}}v = 0 \ \text{ has solution } (\xi, u, v) > 0 \\ &\Leftrightarrow \ 0 \neq \begin{bmatrix} A_{\hat{x}} \\ I_{\hat{x}} \\ -\nabla f(\hat{x}) \end{bmatrix} x \geq 0 \ \text{ has no solution } x \\ &\qquad \text{(by Stiemke's Theorem [9, Theorem 2.4.7])} \\ &\Leftrightarrow \ \operatorname{conj} \begin{bmatrix} A_{\hat{x}} \\ I_{\hat{x}} \\ -\nabla f(\hat{x}) \end{bmatrix} \subseteq \ker \begin{bmatrix} A_{\hat{x}} \\ I_{\hat{x}} \\ -\nabla f(\hat{x}) \end{bmatrix} \\ &\Leftrightarrow \ \operatorname{conj} \begin{bmatrix} A_{\hat{x}} \\ I_{\hat{x}} \\ -\nabla f(\hat{x}) \end{bmatrix} \subseteq \ker H, \end{aligned}$$

the last inclusion following from (4). Hence (iii) holds with $\bar{x}$ replaced with $\hat{x}$, which means that (i) holds with $\bar{x}$ replaced by $\hat{x}$, and since $\nabla f(\bar{x}) = \nabla f(\hat{x})$ by Theorem 1, the required result follows. □

Example 7. The example

$$\begin{array}{ll} \text{minimize} & (x_1-1)^2+(x_2+1)^2 \\ \text{subject to} & x_1, x_2 \geqslant 0 \end{array}$$

does not have the MPS property at its unique solution point $\bar{x}=(1,0)$. However, $\bar{x}$ is nondegenerate and does not satisfy (iii). This shows that condition (4) cannot be removed for convex quadratic programs as is the case for monotone linear complementarity problems (see Theorem 13).

We proceed now to show how the MPS property leads to an exact solution of a convex program by minimizing objective function in a sufficiently small neighborhood of any solution point. For that purpose we need a strong upper-semicontinuity result for linear programs due to Polyak and Tretiyakov [18, Lemma 4], for which we give a simple derivation. Robinson [19, Lemma 3.5] gives a stronger version of this result. We employ an extreme point characterization of a possibly unbounded solution set of a linear program which is based on a Goldman–Tucker characterization [8, Theorem 15] of such a set. Our characterization is in terms of vertices of polyhedral sets (S and T) which depend on the feasible region but not the objective function.

Lemma 8. *Let*

$$S:=\{x \mid Ax \geqslant b, x \geqslant 0\}, \qquad \bar{S}=\underset{x\in S}{\operatorname{argmin}}\langle c, x\rangle \neq \emptyset, \tag{9}$$

where $A \in \mathbb{R}^m$ *and* $c \in \mathbb{R}^n$. *Let*

$$T:=\{x \mid Ax \geqslant 0, x \geqslant 0, \langle e, x\rangle = 1\}.$$

Then

$$\bar{S}=\{x \mid x=rU+sV, r \geqslant 0, \langle e, r\rangle =1, s\geqslant 0\} \tag{10}$$

where

$$U:=\begin{bmatrix} U_1 \\ \vdots \\ U_l \end{bmatrix} = \text{arg vertex} \min_{x\in S}\langle c, x\rangle \subseteq \text{arg vertex } S \tag{11}$$

and

$$V:=\begin{bmatrix} V_1 \\ \vdots \\ V_m \end{bmatrix} = \begin{cases} \emptyset & \textit{if } T=\emptyset \textit{ or } \min_{x\in T}\langle c, x\rangle > 0, \\ \text{arg vertex} \min_{x\in T}\langle c, x\rangle, & \textit{otherwise,} \end{cases}$$
$$\subseteq \text{arg vertex } T. \tag{12}$$

Proof. Note that either U or V may be empty. By [8, Theorem 15, p. 89], (10) holds with U being defined by (11) and V being the finite set of extreme directions of optimal rays of $\min_{x\in S}\langle c, x\rangle$. But by [8, Lemma 9, p. 87], the rows of V are the extreme points of the set

$$\{x \mid x\in T, \langle c, x\rangle \leqslant 0\}=\{x \mid x\in T, \langle c, x\rangle =0\}$$

the equality following from (9). However, these extreme points are precisely the extreme point solutions of $\min_{x\in T}\langle c, x\rangle$ when $\min_{x\in T}\langle c, x\rangle \le 0$, from which (12) follows. □

We use the above lemma to establish Polyak and Tretiyakov's strong upper-semicontinuity of solution sets of linear programs with a perturbed objective function which is stronger than the upper semicontinuity result of Meyer [16, Theorem 2].

Theorem 9 [18, Lemma 4]. *Let $\{c^k\}\to c$ such that* $\operatorname{argmin}_{x\in S}\langle c^k, x\rangle \neq \emptyset$, *where* $S := \{x \mid Ax \ge b, x \ge 0\}$. *Then*

$$\arg\min_{x\in S}\langle c^k, x\rangle \subseteq \arg\min_{x\in S}\langle c, x\rangle \neq \emptyset \quad \text{for } k\ge K, \text{ some } K. \tag{13}$$

Proof. Let $\bar{S}^k := \operatorname{argmin}_{x\in S}\langle c^k, x\rangle$, then by Lemma 8,

$$\bar{S}^k = \{x \mid x = rU^k + sV^k,\ r\ge 0,\ \langle e, r\rangle = 1,\ s\ge 0\} \tag{14}$$

where U^k and V^k are defined in (11) and (12) with c replaced by c^k. Since $U^k \subseteq$ arg vertex S and $V^k \subseteq$ arg vertex T and the sets arg vertex S and arg vertex T are finite and independent of k, it follows that for $k \ge K$ for some K, there is a fixed finite number, say l, of subsets $\{(U^{k_1}, V^{k_1}), \ldots, (U^{k_l}, V^{k_l})\}$ of $\{(\text{arg vertex } S, \text{arg vertex } T)\}$ that appear infinitely often in the sequence $\{(U^k, V^k)\}$ defining $\bar{S}^k$ in (14). For each (U^{k_j}, V^{k_j}), $j = 1, \ldots, l$, the corresponding $\bar{S}^{k_j}$ defined by (14) solves both $\min_{x\in S}\langle c^k, x\rangle$ for $k \ge K$, and $\min_{x\in S}\langle c, x\rangle$. □

We can immediately use the above theorem to show that for a differentiable objective function minimization over a polyhedral set, the MPS property ensures finite termination at an exact solution for any convergent algorithm which periodically solves $\min_{x\in S}\langle\nabla f(x^k), x\rangle$ where $\{x^k\}$ are the algorithm iterates.

Theorem 10. *Let $f:\mathbb{R}^n\to\mathbb{R}$ be a continuously differentiable function on $\mathbb{R}^n$, let $S := \{x \mid Ax \ge b, x\ge 0\}$, let $\{x^k\}\to\bar{x}$, $\{x^k\}\subseteq S$ and let $\bar{S}^k := \operatorname{argmin}_{x\in S}\langle\nabla f(x^k), x\rangle \neq \emptyset$. Then*

$$\bar{S}^k \subseteq \operatorname*{argmin}_{x\in S}\langle\nabla(\bar{x}), x\rangle \neq \emptyset \quad \text{for } k \ge K, \text{ for some } K. \tag{15}$$

If in addition $\bar{x}\in\operatorname{argmin}_{x\in S} f(x)$ and the MPS property is satisfied, that is

$$\operatorname*{argmin}_{x\in S}\langle\nabla f(\bar{x}), x\rangle \subseteq \operatorname*{argmin}_{x\in S} f(x), \tag{16}$$

then

$$\bar{S}^k \subseteq \operatorname*{argmin}_{x\in S} f(x) \quad \text{for } k\ge K, \text{ for some } K. \tag{17}$$

Proof. The inclusion of (15) follows from (13) of Theorem 9 by setting $c^k = \nabla f(x^k)$ and invoking the continuity of $\nabla f(x)$. The inclusions (15) and (16) imply (17). □

We note in passing that Robinson [19, Lemma 3.5] established the stronger result where $\operatorname{argmin}_{x\in S}\langle c, x\rangle$ in (13) is replaced with $\operatorname{argmin}_{x\in \bar{S}_L}\langle c^k, x\rangle$, where $\bar{S}_L := \operatorname{argmin}_{x\in S}\langle c, x\rangle$. This strengthening can be reflected in Theorem 10 by replacing $\operatorname{argmin}_{x\in S}\langle \nabla f(\bar{x}), x\rangle$ in (15) with $\operatorname{argmin}_{x\in \bar{S}_L}\langle \nabla f(x^k), x\rangle$, where $\bar{S}_L :=$ $\operatorname{argmin}_{x\in S}\langle \nabla f(\bar{x}), x\rangle$.

3. Linear complementary problems

In [13] it was shown that the existence of some nondegenerate solution to a monotone linear complementarity problem was sufficient for the MPS property to hold, as well as for the equivalent convex quadratic program (19) below to have a weak sharp minimum. In this section we obtain the rather surprising result that all these properties are equivalent (Theorem 13). They all lead to the useful property that solving a linearized complementarity problem at a point in a sufficiently small neighborhood of a solution gives an exact solution (Corollary 14). Throughout this section we consider the linear complementarity problem

$$Mx + q \geqslant 0, \quad x \geqslant 0, \quad \langle x, Mx + q\rangle = 0, \tag{18}$$

where $M \in \mathbb{R}^{n\times n}$ and $q \in \mathbb{R}^n$. We define the equivalent quadratic program

$$0 = \min_{x\in S} f(x) \tag{19}$$

where

$$S := \{x \mid Mx + q \geqslant 0, x \geqslant 0\}, \quad f(x) := \langle x, Mx + q\rangle \tag{20}$$

and the solution set of (18) and (19) as

$$\bar{S} := \{x \mid x \in S, f(x) = 0\}. \tag{21}$$

We say that the linear complementarity problem in *nondegenerate* if it possesses a nondegenerate solution, that is

$$\hat{x} \in \bar{S}, \quad \hat{x} + M\hat{x} + q > 0.$$

We begin with an elementary result which shows that for any feasible linear complementarity problem, not necessarily monotone, the linearized problem is always solvable. This result also follows from [20, Lemma 3.2.4].

Lemma 11. $S \neq \emptyset \Rightarrow \operatorname{argmin}_{y\in S}\langle \nabla f(x), y\rangle \neq \emptyset$ *for all* $x \in S$.

Proof. The dual of the feasible primal linear program $\min_{y\in S}\langle \nabla f(x), y\rangle$ is

$$\begin{aligned} &\text{maximize} \quad \langle -q, u\rangle \\ &\text{subject to} \quad M^{\mathrm{T}}u \leqslant (M + M^{\mathrm{T}})x + q, \\ &\qquad\qquad\quad u \geqslant 0, \end{aligned}$$

for which $u=x\in S$ is feasible. Hence the feasible primal linear program is solvable. □

We next show that the linearized complementarity problem at any point in a sufficiently small neighborhood of any *nondegenerate vertex solution* is uniquely solved by the nondegenerate vertex solution. Note that the linear complementarity problem need not be monotone, and the point at which the linearization is made could be infeasible.

Theorem 12. *Let $\bar{x}$ be a nondegenerate vertex solution of the LCP* (18). *Then $\bar{x}$ is a locally unique solution of* (18) *and there exists a ball $\mathbb{B}_\delta(\bar{x})$ such that $\bar{x}$ is the unique solution of the linear program*

$$\min_{x\in S}\langle\nabla f(y), x\rangle \quad \forall y\in\mathbb{B}_\delta(\bar{x}). \tag{22}$$

Proof. That $\bar{x}$ is a locally unique solution of (18) follows from [10, Corollary 3.2]. By Theorem 1 we have that $\bar{x}\in\operatorname{argmin}_{x\in S}\langle\nabla f(\bar{x}), x\rangle$. We also have that $\bar{u}=\bar{x}$ solves the dual linear program

$$\begin{aligned}&\underset{(u,v)}{\text{maximize}} && \langle -q, u\rangle\\ &\text{subject to} && v=-M^{\mathrm{T}}u+\nabla f(\bar{x}),\\ &&& (u,v)\geq 0.\end{aligned}$$

Hence optimal dual basic variables (or optimal reduced costs) are given by

$$\bar{v}_I = M_I\bar{x}+q_I>0, \qquad \bar{u}_J=\bar{x}_J>0,$$

where $I\cup J=\{1,\dots,n\}$ and $I\cap J=\emptyset$, by nondegeneracy. It follows from [11, Theorem 2.1] that $\bar{x}$ is the unique solution of $\min_{x\in S}\langle\nabla f(\bar{x}), x\rangle$, and by [12, Theorem 4] it follows that $\bar{x}=\min_{x\in S}\langle\nabla f(y), x\rangle$ for all y satisfying $\|\nabla f(y)-\nabla f(\bar{x})\|\leq\varepsilon$ for some $\varepsilon>0$. The desired conclusion of the theorem follows by letting $0<\delta\leq \varepsilon/(\|M+M^{\mathrm{T}}\|)$. □

A practical consequence of Theorem 12 is that for a nonmonotone LCP with some nondegenerate vertex solution it is advisable to periodically solve the linear program given by (22), no matter what algorithm one is using, because $\mathbb{B}_\delta(\bar{x})$ may contain a current iterate, and hence an exact solution could be obtained by solving the linear program.

We now establish the principal result of this section, namely the equivalence of the MPS property, the existence of a weak sharp minimum, nondegeneracy and a normal cone inclusion property. Note that the nondegeneracy assumption for the LCP is equivalent to the nondegeneracy assumption of Dunn [5] specialized to quadratic programs, and also employed by Calamai and Moré [4].

Theorem 13. *Let the LCP* (18) *be monotone, that is let M be positive semidefinite and let $S \neq \emptyset$. Then $\bar{S} \neq \emptyset$ and the following are equivalent:*

(i)–(iv) *as in Theorem* 6 *with S, $\bar{S}$ and f defined by* (20) *and* (21).

(v) *The LCP* (18) *is nondegenerate.*

Proof. The implication $S \neq \emptyset \Rightarrow \bar{S} \neq \emptyset$ is standard for M positive semidefinite. The equivalence of (i)–(iv) follows from Theorem 6. The implication (v)⇒(i) follows (without the extra assumption used in Theorem 6) from [13, Lemma 2.2]. We now establish the implication (i)⇒(v) by contradiction. Note that the same implication from Theorem 6 cannot be used directly because it applies to $\min_{x \in S}\langle x, Mx+q\rangle$ and its dual and not to the LCP (18).

Let $\bar{x} \in \bar{S}$, $\operatorname{argmin}_{x \in S}\langle \nabla f(\bar{x}), x\rangle \subseteq \bar{S}$ and suppose that (18) is *degenerate*, that is, it is not nondegenerate. Then, using (i),

$$\min_{x,\,\varepsilon}\{-\varepsilon \mid Mx+q \geqslant 0, x \geqslant 0, \langle \nabla f(\bar{x}), x-\bar{x}\rangle \leqslant 0, (I+M)x+q \geqslant \varepsilon e\} = 0. \quad (23)$$

The dual of the linear program (23) is

$$\max_{(u,\,v,\,\xi) \geqslant 0} \{-\langle q, u+v\rangle - \langle \bar{x}, M\bar{x}\rangle \xi \mid M^{\mathrm{T}}u + (I+M^{\mathrm{T}})v - (q+(M+M^{\mathrm{T}})\bar{x})\xi \leqslant 0, \langle e, v\rangle = 1\} = 0.$$

Let (u, v, ξ) be optimal dual variables, then setting the optimal dual objective equal to zero and premultiplying the first dual constraint by $(u+v)$ gives:

$$\langle q, u+v\rangle = -\langle \bar{x}, M\bar{x}\rangle \xi, \quad \langle u+v, M^{\mathrm{T}}u + (I+M^{\mathrm{T}})v - \xi q - \xi(M+M^{\mathrm{T}})\bar{x}\rangle \leqslant 0,$$

$$\langle e, v\rangle = 1, \quad (u, v, \xi) \geqslant 0$$

$$\Rightarrow \langle u+v, v\rangle + \langle u+v, M^{\mathrm{T}}(u+v)\rangle + \xi^2\langle \bar{x}, M\bar{x}\rangle - \xi\langle u+v, (M+M^{\mathrm{T}})\bar{x}\rangle \leqslant 0,$$

$$\langle e, v\rangle = 1, \quad (u, v, \xi) \geqslant 0$$

$$\Rightarrow 0 \leqslant \langle u+v, v\rangle + (u+v-\xi\bar{x}, M(u+v-\xi\bar{x})\rangle \leqslant 0,$$

$$\langle e, v\rangle = 1, \quad (u, v, \xi) \geqslant 0$$

$$\Rightarrow v = 0, \quad \langle e, v\rangle = 1,$$

which is a contradiction. □

The following corollary follows from the above theorem and Theorem 10.

Corollary 14. *Let the assumptions of Theorem* 13 *hold together with one of the conditions* (i)–(v). *Let $\{x^k\} \to \bar{x} \in \bar{S}$, $\{x^k\} \subseteq S$ and let $\bar{S}^k := \operatorname{argmin}_{x \in S}\langle \nabla f(x^k), x\rangle$. Then*

$$\bar{S}^k \subseteq \bar{S} \quad \textit{for } k \geqslant K, \textit{ for some } K. \quad \square$$

In fact a stronger result can be easily established with the help of Theorems 9 and 13, which we merely state: For a nondegenerate monotone linear complementarity problem, there exists a $\delta > 0$ such that

$$\arg\min_{y \in S}\langle \nabla f(x), y\rangle \subseteq \bar{S} \quad \forall x \in S \cap [\bar{S} + \mathbb{B}_\delta(0)].$$

References

[1] F.A. Al-Khayyal and J. Kyparisis, "Finite convergence of algorithms for nonlinear programs and variational inequalities," to appear in: *Journal of Optimization Theory and Applications.*

[2] D.P. Bertsekas, "Necessary and sufficient conditions for a penalty method to be exact," *Mathematical Programming* 9 (1975) 8-99.

[3] J.V. Burke and M.C. Ferris, "The sharpness of functions on sets," in preparation (1991).

[4] P.H. Calamai and J.J. Moré, "Projected gradient methods for linearly constrained problems," *Mathematical Programming* 39 (1987) 93-116.

[5] J.C. Dunn, "On the convergence of projected gradient processes to singular critical points," *Journal of Optimization Theory and Applications* 55 (1987) 203-216.

[6] M.C. Ferris, "Finite termination of the proximal point algorithm," *Mathematical Programming* 50 (1991) 359-366.

[7] M.C. Ferris, "Weak sharp minima and penalty functions in mathematical programming," Technical Report 779, Computer Sciences Department, University of Wisconsin (Madison, WI, 1988).

[8] A.J. Goldman and A.W. Tucker, "Theory of linear programming," in: H.W. Kuhn, and A.W. Tucker, eds., *Linear Inequalities and Related Systems* (Princeton University Press, Princeton, NJ. 1956) pp. 53-97.

[9] O.L. Mangasarian, *Nonlinear Programming* (McGraw-Hill, New York, 1969).

[10] O.L. Mangasarian, "Locally unique solutions of quadratic programs, linear and nonlinear complementarity problems,"*Mathematical Programming* 19 (1980) 200-212.

[11] O.L. Mangasarian, "Optimal simplex tableau characterization of unique and bounded solutions of linear programs," *Journal of Optimization Theory and Applications* 35 (1981) 123-128.

[12] O.L. Mangasarian, "A stable theorem of the alternative: an extension of the Gordan theorem," *Linear Algebra and its Applications* 41 (1981) 209-223.

[13] O.L. Mangasarian, "Error bounds for nondegenerate monotone linear complementarity problems," *Mathematical Programming (Series B)* 48 (1990) 437-445.

[14] O.L. Mangasarian, "A simple characterization of solution sets of convex programs," *Operations Research Letters* 7(1) (1988) 21-26.

[15] O.L. Mangasarian and R.R. Meyer, "Nonlinear perturbation of linear programs," *SIAM Journal on Control and Optimization* 17(6) (1979) 745-752.

[16] R.R. Meyer, "Continuity properties of linear programs," Technical Report 373, Computer Sciences Department, University of Wisconsin (Madison, WI, 1979).

[17] B.T. Polyak, *Introduction to Optimization* (Optimization Software, Inc., Publications Division, New York, 1987).

[18] B.T. Polyak and N.V. Tretiyakov, "Concerning an iterative method for linear programming and its economic interpretation," *Economics and Mathematical Methods* 8(5) (1972) 740-751. [English translation: *Matekon* 10(3) (1974) 81-100.]

[19] S. M. Robinson, "Local structure of feasible sets in nonlinear programming, Part II: Nondegeneracy," *Mathematical Programming Study* 22 (1984) 217-230.

[20] T.-H. Shiau, "Iterative linear programming for linear complementarity and related problems," Technical Report 507, Computer Sciences Department, University of Wisconsin (Madison, WI, 1983).