Mathematical Programming manuscript No.
(will be inserted by the editor)

Stephen J. Wright*

On Reduced Convex QP Formulations of Monotone LCPs

February 20, 2001

Abstract. Techniques for transforming convex quadratic programs (QPs) into monotone linear complementarity problems (LCPs) and vice versa are well known. We describe a class of LCPs for which a reduced QP formulation—one that has fewer constraints than the "standard" QP formulation—is available. We mention several instances of this class, including the known case in which the coefficient matrix in the LCP is symmetric.

Key words. Monotone Linear Complementarity Problems, Convex Quadratic Programming, Karush-Kuhn-Tucker Conditions

1. Introduction

In this note, we consider linear complementarity problems (LCPs) and convex quadratic programs (QPs) over closed convex cones, and examine the relationship between LCP and QP formulations of the same problem. We show that for a subclass of LCPs, it is possible to define a QP that has fewer constraints than the standard QP reformulation and whose primal-dual solution yields a solution of the LCP.

Our work is related to that of Robinson [7,8], who discusses methods for reducing variational inequalities (possibly nonlinear and nonmonotone) that have certain structural properties. There is a subclass of problems for which the reduction techniques discussed here and those of Robinson are identical. We mention some problems of this type in Section 4 and discuss the relationship between the reduction techniques in more detail there.

The significance of our results derives partly from the fact that software for solving QP is generally more prevalent than software for LCP. Given some LCP formulation of a problem, and a code for solving QP, it is often to our advantage to find the most compact QP representation of the problem possible before calling the code to solve it.

Stephen J. Wright: Mathematics and Computer Science Division, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois 60439; `wright@mcs.anl.gov`

* Research supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Advanced Scientific Computing Research, U.S. Department of Energy, under Contract W-31-109-Eng-38.

Section 2 provides some background and presents an existence result for the solution of monotone LCP. Section 3 proves the main results about reduced QP formulations. Some examples are given in Section 4.

2. Background

We now define affine variational inequalities, linear complementarity problems, and quadratic programming problems over a closed convex cone in Euclidean space R^n, and outline the techniques by which a given problem can be formulated by any one of these techniques. We also state results about the equivalence of these formulations and existence of solutions.

We use $\langle \cdot, \cdot \rangle$ to denote an inner product in R^n. (All our examples use $\langle x, y \rangle = x^T y$.) We use M, Q, R, S, and A to denote linear operators on R^n (or, equivalently, their matrix representations), and M^*, Q^*, etc., to denote their adjoints.

Closed convex cones are closed sets $K \subset \mathsf{R}^n$ such that for any vectors $x \in K$ and $y \in K$, we have that $\alpha x + \beta y \in K$ for all $\alpha \geq 0$ and $\beta \geq 0$. Note in particular that $0 \in K$, that by setting $\alpha \in [0,1]$ and $\beta = 1 - \alpha$ we verify the convexity property, and that $x + K \subset K$ for all $x \in K$. The cone K is said to be *polyhedral* if there exist a finite set of vectors $\{a_1, a_2, \ldots, a_L\}$ such that

$$K = \left\{ x \in \mathsf{R}^n \mid x = \sum_{l=1}^{L} \gamma_l a_l, \; \gamma_l \geq 0 \text{ for all } l \right\}. \tag{1}$$

The *polar cone* for K is defined by

$$K^\circ \stackrel{\text{def}}{=} \{ s \mid \langle y, s \rangle \leq 0 \text{ for all } y \in K \}.$$

It is easy to verify that $(K^\circ)^\circ = K$; see Rockafellar [9, p. 121]. When K is polyhedral with the form (1), K° is also polyhedral and is given explicitly as

$$K^\circ = \{ s \mid \langle s, a_l \rangle \leq 0 \;\text{ for all } l = 1, 2, \ldots, L \}$$

(see Rockafellar [9, p. 122]). The *normal cone* for K at a point x is defined by

$$N_K(x) \stackrel{\text{def}}{=} \begin{cases} \{ s \mid \langle s, y - x \rangle \leq 0 \text{ for all } y \in K \}, & \text{if } x \in K, \\ \emptyset & \text{if } x \notin K. \end{cases} \tag{2}$$

It follows immediately that $K^\circ = N_K(0)$.

Given a convex function f on R^n, we define the conjugate function f^* by

$$f^*(y) = \sup_x \{ \langle x, y \rangle - f(x) \mid x \in \text{ri}(\text{dom} f) \}.$$

The subgradient of f is the multifunction defined by

$$\partial f(x) = \{ y \mid f(z) \geq f(x) + \langle y, z - x \rangle \;\text{ for all } z \}.$$

As in Robinson [7], we use these definitions to note the following relationship:

$$y \in N_K(x) \;\Leftrightarrow\; x \in N_{K^\circ}(y). \tag{3}$$

A proof of this claim follows if we note that

$$N_K(x) = \partial I_K(x),$$

where I_K is the indicator function for K (which takes on the value 0 on K and ∞ otherwise), use the fact that $I_K^* = I_{K^\circ}$ (Rockafellar [9, Theorem 14.1]), and then apply Theorem 23.5 of [9].

Consider the affine variational inequality problem over the closed convex cone $K \subset \mathsf{R}^n$:

$$\text{Find } x \in \mathsf{R}^n \text{ such that } q - Mx \in N_K(x), \tag{4}$$

where $M \in \mathsf{R}^{n\times n}$ and $q \in \mathsf{R}^n$ are given. If x solves (4), we have from $x + y \in K$ for all $y \in K$ and the definition (2) that

$$\langle q - Mx, y\rangle = \langle q - Mx, (x+y) - x\rangle \le 0, \text{ for all } y \in K.$$

Therefore $q - Mx \in K^\circ$, so it is natural to write the LCP associated with (4) as follows:

$$\text{Find } x \in K \text{ such that } q - Mx \in K^\circ, \; \langle x, q - Mx\rangle = 0. \tag{5}$$

In fact, Proposition 1.5.2 of Cottle, Pang, and Stone [1] shows that problems (4) and (5) are equivalent.

We are interested in monotone problems, those for which M is monotone but not necessarily self-adjoint. The formulation (5) reduces to the standard monotone LCP if we take $K = \mathsf{R}^n_+$, the nonnegative orthant. We obtain the mixed monotone LCP if we set $K = \mathsf{R}^{\bar{n}}_+ \times \mathsf{R}^{n-\bar{n}}$ for some $\bar{n}$ strictly between 0 and n. The polar cone in this case is $K^\circ = \mathsf{R}^{\bar{n}}_- \times \{0\}$.

Consider now the quadratic programming problem (QP) over a closed convex cone $L \subset \mathsf{R}^m$:

$$\min_{z\in\mathsf{R}^n} \; \tfrac{1}{2}\langle z, Qz\rangle - \langle c, z\rangle, \text{ subject to } Az - b \in L, \tag{6}$$

where Q is self-adjoint and monotone. The standard technique for reformulating (5) as a quadratic program (6) is to define the inner product to be the objective function and write

$$\min_x \langle x, Mx - q\rangle \text{ subject to } x \in K, \; q - Mx \in K^\circ, \tag{7}$$

To identify (7) with (6), we define $z = x$ and

$$Q = M + M^*, \; c = q, \; Az = (z, -Mz), \; b = (0, -q), \; L = K \times K^\circ,$$

where, as mentioned earlier, M^* is the adjoint of M. Conversely, the standard technique for reformulating (6) in the form (5) is via the Karush-Kuhn-Tucker (KKT) optimality conditions for (6), which are that there exists a vector v such that

$$Qz - c + A^* v = 0, \; Az - b \in L, \; v \in L^\circ, \; \langle v, Az - b\rangle = 0. \tag{8}$$

(These conditions can be derived for example by applying results from Rockafellar [9, Ch. 28], together with elementary observations about L, L°, and normal cones.) Because of the assumed monotonicity of Q and convexity of L, the conditions (8) are sufficient as well as necessary for solving (6).

For the particular QP (7) that arises as a reformulation of the LCP (5), the conditions (8) reduce to the following:

$$(M+M^*)x - q - M^*u + w = 0, \tag{9a}$$
$$(x, -Mx) - (0, -q) \in K \times K^\circ, \tag{9b}$$
$$(w, u) \in K^\circ \times K, \tag{9c}$$
$$\langle (x, q - Mx), (w, u) \rangle = 0. \tag{9d}$$

Note that this LCP is quite different from the original form (5), in that the number of variables is significantly greater. However, an equivalence relationship between these two LCPs and the QP (7) can be derived in the case in which K is in addition polyhedral.

Theorem 1. *Suppose that K is a closed convex cone in R^n, and that (5) is feasible; that is, there exists $x \in K$ such that $q - Mx \in K^\circ$. Then*

(i) if K is polyhedral, the LCP (5) has a solution; and
(ii) any solution of (5) also solves (7), and conversely.

Proof. The proof of (i) follows immediately from Gowda and Seidman [3, Theorem 3.1].

For (ii), we see immediately that any solution of (5) is feasible in (7) with an objective value of 0. Since, as observed earlier, 0 is a lower bound on the objective in (7), we conclude that x solves (7).

It remains to prove only that any solution x of (7) also solves (5). By necessity of the conditions (9), there exists a vector (u, w) such that (9) is satisfied by (x, u, w). We now show that this x solves (5) by using a generalization of the argument of Theorem 3.1.2 of [1]. By taking the inner product of (9a) with $x - u$, we obtain

$$\begin{aligned}
0 &= \langle x-u, Mx - q + M^*(x-u) + w \rangle \\
&\geq \langle x-u, Mx-q+w \rangle && \text{by monotonicity of } M \\
&= \langle x, Mx-q \rangle + \langle (x, q-Mx), (w,u) \rangle - \langle u, w \rangle \\
&= \langle x, Mx-q \rangle - \langle u, w \rangle && \text{by (9d)} \\
&\geq \langle x, Mx-q \rangle && \text{since } u \in K,\ w \in K^\circ.
\end{aligned}$$

Since $x \in K$ and $q - Mx \in K^\circ$, we have also that $\langle x, Mx-q \rangle \geq 0$, so it follows that $\langle x, Mx - q \rangle = 0$ and hence that x solves (5).

Gowda [2] shows that a quadratic functions attains its minimizer over a closed polyhedral convex cone when it is bounded below on the cone; this result can be used to prove existence of the solution to (7) directly whenever the feasible region for this problem is nonempty.

3. Reduced QP Formulations

We now examine special structures of the operator M that allow us to define a QP reformulation of the LCP (5) with possibly fewer constraints than the standard reformulation (7). Our results in this section extend the observation of Cottle, Pang, and Stone [1, Section 1.4]. A slightly generalized version of the latter result states that when M is self-adjoint and monotone, the LCP (5) is equivalent to the following QP:

$$\min_x \tfrac{1}{2}\langle x, Mx\rangle - \langle q, x\rangle \text{ subject to } x \in K. \tag{10}$$

(To verify the equivalence, note that the necessary and sufficient optimality conditions for (10) are $q - Mx \in N_K(x)$, which is equivalent to (5).) By comparing with (7), we see that the quadratic term in the objective of (10) differs and that the constraint $q - Mx \in K^\circ$ is not present.

The special structure of M that we analyze in this section is defined with respect to a subspace T of R^n. A projection onto this subspace is denoted by P_T, where

$$P_T x = \arg\min_{y\in T}\langle y - x, y - x\rangle. \tag{11}$$

Note that P_T is a self-adjoint linear operator and that $P_T P_T = P_T$.

The subspace orthogonal to T is $T^\perp = \{y \mid \langle y, x\rangle = 0 \text{ for all } x \in T\}$. We have that

$$I = P_T + P_{T^\perp}, \tag{12}$$

that is, any vector $x \in \mathsf{R}^n$ can be decomposed as $x = P_T x + P_{T^\perp} x$.

In this section we assume certain properties on two-sided projections of M onto T and its complement $T^\perp$. To be specific, we are interested in M for which there exist operators S, Q, and R on R^n, such that

$$P_T M P_T = P_T S P_T, \;\; S \text{ monotone and self-adjoint}, \tag{13a}$$
$$P_T M P_{T^\perp} = P_T R P_{T^\perp}, \tag{13b}$$
$$P_{T^\perp} M P_T = P_{T^\perp}(-R^*) P_T = -P_{T^\perp} R^* P_T, \tag{13c}$$
$$P_{T^\perp} M P_{T^\perp} = P_{T^\perp} Q P_{T^\perp}, \;\; Q \text{ monotone and self-adjoint}. \tag{13d}$$

A number of identities follow from these properties. For example, we have

$$\begin{aligned} P_T(M + M^*) &= P_T(M + M^*)P_T + P_T(M + M^*)P_{T^\perp} \\ &= P_T M P_T + P_T M^* P_T + P_T M P_{T^\perp} + P_T M^* P_{T^\perp} \\ &= P_T M P_T + (P_T M P_T)^* + P_T M P_{T^\perp} + (P_{T^\perp} M P_T)^* \\ &= 2P_T S P_T + P_T R P_{T^\perp} + (-P_{T^\perp} R^* P_T)^* \\ &= 2P_T S P_T, \end{aligned} \tag{14}$$

where we used the self-adjoint property of S. Similarly

$$P_{T^\perp}(M + M^*) = 2P_{T^\perp} Q P_{T^\perp}. \tag{15}$$

For our problem class of interest, we assume too that T and K are related in a certain way. Defining

$$P_T K \stackrel{\text{def}}{=} \{v \mid v = P_T x \text{ for some } x \in K\}, \tag{16}$$

we assume that

$$P_T K \subset K, \quad P_{T^\perp} K \subset K. \tag{17}$$

Similar inclusions for K° follow by a simple argument: Given any $y \in K^\circ$, we have from $P_T K \subset K$ that $\langle y, P_T v\rangle \le 0$ for all $v \in K$. Since $\langle y, P_T v\rangle = \langle P_T y, v\rangle$, we have that $\langle P_T y, v\rangle \le 0$ for all $v \in K$, so that $P_T y \in K^\circ$. We deduce that

$$P_T K^\circ \subset K^\circ, \quad P_{T^\perp} K^\circ \subset K^\circ. \tag{18}$$

We also have the following lemma.

Lemma 1. *Suppose that (17) holds. Then for any $x \in K$, we have*

$$N_{P_T K}(P_T x) = \{v \mid P_T v \in N_K(P_T x)\}. \tag{19}$$

Proof. Note first that $P_T x \in P_T K$, since $x \in K$. Therefore, from the definition (2), we have

$$\begin{aligned}
N_{P_T K}(P_T x) &= \{v \mid \langle v, P_T t - P_T x\rangle \le 0, \text{ all } t \in K\} \\
&= \{v \mid \langle v, P_T t\rangle - \langle v, P_T x\rangle \le 0, \text{ all } t \in K\} \\
&= \{v \mid \langle P_T v, t\rangle - \langle P_T v, P_T x\rangle \le 0, \text{ all } t \in K\} \\
&= \{v \mid \langle P_T v, t - P_T x\rangle \le 0, \text{ all } t \in K\} \\
&= \{v \mid P_T v \in N_K(P_T x)\}.
\end{aligned}$$

Similar relationships follow from (17) and (18); in particular, for any $y \in K^\circ$, we have

$$N_{P_{T^\perp} K^\circ}(P_{T^\perp} y) = \{u \mid P_{T^\perp} u \in N_{K^\circ}(P_{T^\perp} y)\}. \tag{20}$$

The following technical lemma is also useful in proving our main result.

Lemma 2. *Let x_1, x_2, v_1, and v_2 be vectors such that*

$$x_1 \in K, \; x_2 \in K; \quad v_1 \in N_K(x_1), \; v_2 \in N_K(x_2); \quad \langle v_2, x_1\rangle = \langle v_1, x_2\rangle = 0.$$

Then

$$v_1 + v_2 \in N_K(x_1 + x_2).$$

Proof. Since $v_1 \in N_K(x_1)$ and $v_2 \in N_K(x_2)$, we have that

$$\langle v_1, t - x_1\rangle \le 0, \;\; \langle v_2, t - x_2\rangle \le 0, \;\; \text{for all } t \in K. \tag{21}$$

But given any $t \in K$, we have that

$$\begin{aligned}
&\langle v_1 + v_2, t - (x_1 + x_2)\rangle \\
&= \langle v_1, t - x_1\rangle - \langle v_1, x_2\rangle + \langle v_2, t - x_2\rangle - \langle v_2, x_1\rangle \\
&= \langle v_1, t - x_1\rangle + \langle v_2, t - x_2\rangle \le 0,
\end{aligned}$$

proving the result. The result can also be proved by making use of the following characterization: $N_K(x) = K^\circ \cap \{v \mid \langle v, x\rangle = 0\}$.

We are now ready to derive our main result, which is to show that under the assumptions on M and K made in this section, a solution of (5) can be obtained from the primal-dual solution of the following convex quadratic program:

$$\begin{aligned} \min\ & \tfrac{1}{4}\langle x, (M+M^*)x\rangle - \langle P_T q, x\rangle && \text{(22a)}\\ \text{subject to}\ & P_{T^\perp}(q - Mx) \in P_{T^\perp}K^\circ, && \text{(22b)}\\ & P_T x \in P_T K. && \text{(22c)} \end{aligned}$$

The (necessary and sufficient) optimality conditions for this problem are as follows:

$$\begin{aligned} -P_T q + \tfrac{1}{2}(M+M^*)x - M^* P_{T^\perp} u + P_T v = 0, && \text{(23a)}\\ P_{T^\perp}(q - Mx) \in P_{T^\perp}K^\circ, && \text{(23b)}\\ P_T x \in P_T K, && \text{(23c)}\\ u \in N_{P_{T^\perp}K^\circ}(P_{T^\perp}(q - Mx)), && \text{(23d)}\\ v \in N_{P_T K}(P_T x). && \text{(23e)} \end{aligned}$$

Because of (19) and (20), we have that

$$P_{T^\perp} u \in N_{K^\circ}(P_{T^\perp}(q - Mx)), \qquad P_T v \in N_K(P_T x).$$

We can therefore rewrite (23) as follows:

$$\begin{aligned} -P_T q + \tfrac{1}{2}(M+M^*)x - M^* P_{T^\perp} u + P_T v = 0, && \text{(24a)}\\ P_{T^\perp}(q - Mx) \in P_{T^\perp}K^\circ, && \text{(24b)}\\ P_T x \in P_T K, && \text{(24c)}\\ P_{T^\perp} u \in N_{K^\circ}(P_{T^\perp}(q - Mx)), && \text{(24d)}\\ P_T v \in N_K(P_T x). && \text{(24e)} \end{aligned}$$

We now show that the primal-dual solution of (22) yields a solution of (4) (equivalently, (5)). By operating on (24a) with $P_{T^\perp}$, we obtain from (13d), the self-adjointness of Q and $P_{T^\perp}$, and the identity (15) that

$$\begin{aligned} 0 &= \tfrac{1}{2}P_{T^\perp}(M+M^*)x - P_{T^\perp}M^*P_{T^\perp}u\\ &= P_{T^\perp}QP_{T^\perp}x - [P_{T^\perp}QP_{T^\perp}]^* u\\ &= P_{T^\perp}QP_{T^\perp}x - P_{T^\perp}QP_{T^\perp}u. \end{aligned} \tag{25}$$

From (24d), and using (3), we obtain

$$P_{T^\perp}(q - Mx) \in N_K(P_{T^\perp}u). \tag{26}$$

By expanding $P_{T^\perp}(q - Mx)$ and using (13) and (25), we obtain

$$\begin{aligned} P_{T^\perp}(q - Mx) &= P_{T^\perp}q - P_{T^\perp}MP_{T^\perp}x - P_{T^\perp}MP_T x\\ &= P_{T^\perp}q - P_{T^\perp}QP_{T^\perp}x + P_{T^\perp}R^*P_T x\\ &= P_{T^\perp}q - P_{T^\perp}QP_{T^\perp}u + P_{T^\perp}R^*P_T x, \end{aligned}$$

so from (26) we have

$$P_{T^\perp} q - P_{T^\perp} Q P_{T^\perp} u + P_{T^\perp} R^* P_T x \in N_K(P_{T^\perp} u). \tag{27}$$

We now operate on (24a) with P_T and use (13), (14), and self-adjointness of P_T and $P_{T^\perp}$ to obtain

$$\begin{aligned} 0 &= -P_T q + P_T S P_T x - P_T M^* P_{T^\perp} u + P_T v \\ &= -P_T q + P_T S P_T x - [P_{T^\perp} M P_T]^* u + P_T v \\ &= -P_T q + P_T S P_T x + [P_{T^\perp} R^* P_T]^* u + P_T v \\ &= -P_T q + P_T S P_T x + P_T R P_{T^\perp} u + P_T v. \end{aligned} \tag{28}$$

Hence, by substitution into (24e), we obtain

$$P_T q - P_T S P_T x - P_T R P_{T^\perp} u \in N_K(P_T x). \tag{29}$$

From Lemma 2, we have by combining (27) and (29) that

$$\begin{aligned} &(P_{T^\perp} q - P_{T^\perp} Q P_{T^\perp} u + P_{T^\perp} R^* P_T x) + (P_T q - P_T S P_T x - P_T R P_{T^\perp} u) \\ &\in N_K(P_{T^\perp} u + P_T x), \end{aligned}$$

so that, using (12) and (13), we have

$$q - P_T M P_T x - P_T M P_{T^\perp} u - P_{T^\perp} M P_{T^\perp} u - P_{T^\perp} M P_T x \in N_K(P_{T^\perp} u + P_T x).$$

If we define

$$x^* = P_{T^\perp} u + P_T x, \tag{30}$$

we see immediately that

$$q - M x^* \in N_K(x^*). \tag{31}$$

We conclude that from the primal-dual solution of (22), we can construct a solution of (4), and therefore of (5). This result, slightly enhanced, can be stated formally as follows.

Theorem 2. *Suppose that for the matrix M, the subspace T, and the closed convex cone K the conditions (13) and (17) (and therefore (18)) are satisfied. Then if (x, u, v) is a primal-dual solution of (22), we have that x^* defined by (30) is a solution of (5). Moreover, if Q in (13d) is strictly monotone on the subspace $T^\perp$—that is, $\langle v, Qv \rangle > 0$ for all $0 \neq v \in T^\perp$—then the primal solution x of (22) also solves (5).*

Proof. We have proved the first statement already in the paragraphs above. For the second statement we have, by taking inner products of (25) with $(x - u)$, that

$$\langle P_{T^\perp}(x - u), Q P_{T^\perp}(x - u) \rangle = 0,$$

so from the strict monotonicity property we have $P_{T^\perp} u = P_{T^\perp} x$. Therefore we can replace $P_{T^\perp} u$ by $P_{T^\perp} x$ in (30), giving the result.

A similar result can be proved if we replace (22) by its dual, by interchanging the roles of T and $T^\perp$. We obtain the following QP:

$$\min \tfrac{1}{4}\langle x, (M+M^*)x\rangle - \langle P_{T^\perp} q, x\rangle \tag{32a}$$
$$\text{subject to } \; P_T(q - Mx) \in P_T K^\circ, \tag{32b}$$
$$P_{T^\perp} x \in P_{T^\perp} K. \tag{32c}$$

We show by similar logic to the analysis of (22) that a primal-dual solution of (32) yields a solution of (5) under the assumptions of this section. The formal result is as follows

Theorem 3. *Suppose that for the matrix M, the subspace T, and the closed convex cone K the conditions (13) and (17) (and therefore (18)) are satisfied. Then if (x,u,v) is a primal-dual solution of (32), we have that x^* defined by*

$$x^* = P_T u + P_{T^\perp} x, \tag{33}$$

is a solution of (5). Moreover, if S is strictly monotone on the subspace T, then the primal solution x of (32) also solves (5).

The significance of Theorems 2 and 3 is that the number of linear equalities and inequalities required to express the relations $P_{T^\perp}(q - Mx) \in P_{T^\perp}K^\circ$, $P_T x \in P_T K$, and so on is often fewer than the corresponding number required to represent $q - Mx \in K^\circ$, $x \in K$ in the standard formulation (6). Therefore, if we have available software for solving convex QPs, we might expect more efficient practical performance from applying it to the formulations (22) and (32) than to (6).

4. Examples

We now consider some examples of problems of the type analyzed in Section 3, illustrating the reduced QP formulations in each case.

Example 1. We believe that most practical instances of the problem structure analyzed in this paper will have the following form. Let the cone $K \subset \mathsf{R}^n$ be a Cartesian product of the form

$$K = K_0 \times K_1, \tag{34}$$

where $K_0 \subset \mathsf{R}^{n_0}$ and $K_1 \subset \mathsf{R}^{n_1}$ are both closed convex cones, with $n = n_0 + n_1$. Assume too that the coefficient matrix M can be written in the form

$$M = \begin{bmatrix} \bar{S} & \bar{R} \\ -\bar{R}^T & \bar{Q} \end{bmatrix}, \tag{35}$$

where $\bar{S} \in \mathsf{R}^{n_0 \times n_0}$ and $\bar{Q} \in \mathsf{R}^{n_1 \times n_1}$ are symmetric positive semidefinite. The vector q and the vector of unknowns x are partitioned correspondingly as follows:

$$\begin{bmatrix} q_0 \\ q_1 \end{bmatrix}, \quad \begin{bmatrix} x_0 \\ x_1 \end{bmatrix}, \quad \text{where } x_0, q_0 \in \mathsf{R}^{n_0}, \; x_1, q_1 \in \mathsf{R}^{n_1}.$$

We now define

$$T = \mathsf{R}^{n_0} \times \{0\}, \quad T^{\perp} = \{0\} \times \mathsf{R}^{n_1}, \tag{36}$$

and note that (17) obviously holds, since

$$P_T K = K_0 \times \{0\}, \quad P_{T^\perp} K = \{0\} \times K_1.$$

We identify the components in (35) with the quantities S, R, and Q from (13) by defining

$$S = \begin{bmatrix} \bar{S} & 0 \\ 0 & 0 \end{bmatrix}, \quad R = \begin{bmatrix} 0 & \bar{R} \\ 0 & 0 \end{bmatrix}, \quad Q = \begin{bmatrix} 0 & 0 \\ 0 & \bar{Q} \end{bmatrix}. \tag{37}$$

By referring to (22), we can write the reduced QP formulation of this mixed monotone LCP as follows:

$$\min_{x_0, x_1} \tfrac{1}{2}(x_0^T \bar{S} x_0 + x_1^T \bar{Q} x_1) - q_0^T x_0, \tag{38a}$$

$$\text{subject to } q_1 + \bar{R}^T x_0 - \bar{Q} x_1 \in K_1^\circ, \tag{38b}$$

$$x_0 \in K_0. \tag{38c}$$

Note that we have modified the formulation (22) by omitting the constraints in which both sides are identically zero. The standard QP formulation (6) would have $2n$ constraints, in contrast to the n constraints needed in (38). The alternative formulation (32) becomes

$$\min_{x_0, x_1} \tfrac{1}{2}(x_0^T \bar{S} x_0 + x_1^T \bar{Q} x_1) - q_1^T x_1, \tag{39a}$$

$$\text{subject to } q_0 - \bar{S} x_0 - \bar{R} x_1 \in K_0^\circ \tag{39b}$$

$$x_1 \in K_1. \tag{39c}$$

Example 1A. If there is rank deficiency in the matrix $\bar{Q}$, the vector x_1 in formulation (38) can be replaced by a lower-dimensional object. In the extreme case of $\bar{Q} = 0$, x_1 does not appear at all. The reduced formulation (38) reduces further to

$$\min_{x_0} \tfrac{1}{2} x_0^T \bar{S} x_0 - q_0^T x_0, \tag{40a}$$

$$\text{subject to } q_1 + \bar{R}^T x_0 \in K_1^\circ, \tag{40b}$$

$$x_0 \in K_0. \tag{40c}$$

This case is covered by the analysis of Robinson [7, Proposition 2]. We can identify the optimality conditions for (40) with Robinson [7, eq. (8)] by defining $-d(\cdot)$ appropriately and setting $Y = K_1$ and $P = K_0$.

If instead we have that $\bar{S} = 0$, then (39) can be used to obtain a reduced problem in which only the variables x_1 appear.

Example 1B. Suppose that $n_1 = 0$, so that $\bar{R}$, $\bar{Q}$, and q_1 are all vacuous. Then (38) reduces to

$$\min_{x_0} \tfrac{1}{2} x_0^T \bar{S} x_0 - q_0^T x_0, \quad \text{subject to } x_0 \in K,$$

where $K = K_0$. This is simply the form (10) whose equivalence to (5) in the case of symmetric positive semidefinite $\bar{S}$ was essentially noted by Cottle, Pang, and Stone [1, Section 1.4]. Again, the reduction of Robinson [7, eq. (8)] yields the same result.

The following alternative, generally less useful formulation is available from (39):

$$\min_{x_0} \tfrac{1}{2} x_0^T \bar{S} x_0, \quad \text{subject to } q_0 - \bar{S} x_0 \in K^\circ.$$

Example 1C. A further special case of Example 1 is the linear programming problem in standard form. Here we have

$$\bar{S} = 0, \;\; \bar{Q} = 0, \;\; \bar{R} = -A^T,$$

with the coordinate cones are defined as

$$K_0 = \mathsf{R}_+^{n_0}, \qquad K_1 = \mathsf{R}^{n_1}.$$

The resulting LCP (5) is then

$$q_1 - Ax_0 = 0, \;\; q_0 + A^T x_1 \le 0, \;\; x_0 \ge 0, \;\; x_0^T (A^T x_1 + q_0) + x_1^T (-Ax_0 + q_1) = 0,$$

which by simple manipulation becomes

$$Ax_0 = q_1, \;\; A^T x_1 \le -q_0, \;\; x_0 \ge 0, \;\; x_0^T q_0 + x_1^T q_1 = 0. \tag{41}$$

The reduced QP form (22) (equivalently (38)) is

$$\min_{z,w} -q_0^T x_0 \;\text{ s.t. } Ax_0 = q_1, \; x_0 \ge 0, \tag{42}$$

which is simply the linear programming problem in standard form. The alternative reduced QP form (32) (equivalently (39)) is

$$\min \; -q_1^T x_1 \;\text{ s.t. } A^T x_1 \le -q_0, \tag{43}$$

which is just the dual of the standard form. In practice, it is usually beneficial to apply software to either (42) or (43), rather than to the larger self-dual form that would arise from the standard QP formulation (6), namely,

$$\begin{aligned} \min \; -q_0^T x_0 - q_1^T x_1 & \\ \text{s.t. } \; Ax_0 &= q_1, \\ A^T x_1 &\le -q_0, \\ x_0 &\ge 0. \end{aligned}$$

Application of linear programming software to this form would be efficient only if the code were able to recognize and exploit the self-dual structure.

Example 2. We now consider the extended linear-quadratic programming (ELQP) problem proposed by Rockafellar [10,11]. Given nonempty polyhedral convex sets $Y \subset \mathsf{R}^{n_0}$ and $Z \subset \mathsf{R}^{n_1}$, matrices $\bar{S}$ and $\bar{Q}$, and vectors q_0 and q_1 with the same form as in Example 1, and a matrix $A \in \mathsf{R}^{n_0 \times n_1}$, the ELQP problem is as follows:

$$\min_{y \in Y} -\langle q_0, y\rangle + \tfrac{1}{2}\langle y, \bar{S}y\rangle + \theta_{Z,\bar{Q}}(q_1 + A^T y), \tag{44}$$

where

$$\theta_{Z,\bar{Q}}(u) = \sup_{z \in Z} \langle z, u\rangle - \tfrac{1}{2}\langle z, \bar{Q}z\rangle. \tag{45}$$

The dual of this problem is

$$\max_{z \in Z} \langle q_1, z\rangle - \tfrac{1}{2}\langle z, \bar{Q}z\rangle - \theta_{Y,\bar{S}}(q_0 - Az), \tag{46}$$

where

$$\theta_{Y,\bar{S}}(v) = \sup_{y \in Y} \langle y, v\rangle - \tfrac{1}{2}\langle y, \bar{S}y\rangle. \tag{47}$$

ELQP has proved to be a highly versatile framework that includes many piecewise linear and piecewise quadratic problems. We consider here the case in which Y and Z are closed convex cones. This subset of ELQP includes linear and quadratic programming problems as special cases. For instance, the constraint $q_1 + A^T y \le 0$ can be enforced by setting $\bar{Q} = 0$ and $Z = \mathsf{R}^{n_1}_+$ in (45). The framework can also incorporate "soft constraints," a modeling technique that is frequently used in practice. In this technique, a violation of a desired inequality is not forbidden, but is discouraged by the inclusion of a quadratic term in the violation in the objective. For instance, if we set

$$Z = \mathsf{R}^{n_1}_+, \qquad \bar{Q} = (\sigma/2)I,$$

for some $\sigma > 0$, then from (45) we have

$$\theta_{Z,\bar{Q}}(q_1 + A^T y) = \frac{1}{2\sigma} \left\| (q_1 + A^T y)_+ \right\|_2^2, \tag{48}$$

where the subscript "+" denotes projection onto $\mathsf{R}^{n_1}_+$.

It is easy to show that the optimality conditions for (44), (45) simply have the form of the LCP in Example 1. These conditions are

$$\begin{aligned} q_0 - \bar{S}y - Az &\in N_Y(y), \\ q_1 + A^T y - \bar{Q}z &\in N_Z(z). \end{aligned}$$

As in Example 1, we have that the reduced QP formulation (22) is

$$\min_{y,z} \tfrac{1}{2}(y^T \bar{S}y + z^T \bar{Q}z) - q_0^T y, \tag{49a}$$

$$\text{subject to } q_1 + A^T y - \bar{Q}z \in Z^*, \tag{49b}$$

$$y \in Y. \tag{49c}$$

The alternative formulation, corresponding to (32), is

$$\min_{y,z} \tfrac{1}{2}(y^T \bar{S}y + z^T \bar{Q}z) - q_1^T z, \tag{50a}$$

$$\text{subject to } q_0 - \bar{S}y - Az \in Y^* \tag{50b}$$

$$z \in Z. \tag{50c}$$

Example 2A. A special case of ELQP is the subproblem that arises in the stabilized sequential quadratic programming (SSQP) method described in Wright [12]. The subproblem to be solved is similar to the one that leads to (48). It has the form

$$\min_{z} \tfrac{1}{2} z^T \bar{Q} z - c^T z + \max_{\lambda \geq 0} \lambda^T (b - Az) - \tfrac{1}{2} \epsilon \|\lambda - \lambda^-\|_2^2, \tag{51}$$

where λ_- is the estimate of λ from the previous iteration and $\epsilon > 0$ is the stabilization parameter. When $\bar{Q}$ is positive semidefinite, this problem has the form of (46), (47) if we set

$$y = \lambda, \quad q_1 = c, \quad q_0 = b + \epsilon \lambda^-, \quad \bar{S} = \epsilon I, \quad Z = \mathsf{R}^n, \quad Y = \mathsf{R}^m_+,$$

and ignore the constant term in the objective. The form (50) is then

$$\min_{z,\lambda} \tfrac{1}{2} \epsilon \|\lambda\|_2^2 + \tfrac{1}{2} z^T \bar{Q} z - c^T z \quad \text{subject to} \quad Az - b + \epsilon(\lambda - \lambda^-) \geq 0,$$

which is equivalent to the form derived by Li and Qi [4, eq. (15)]. We can eliminate λ from this problem (at the cost of some nonsmoothness in the objective) and write it as

$$\min_{z} \frac{1}{2\epsilon} \left\| [b - Az + \epsilon \lambda^-]_+ \right\|_2^2 + \tfrac{1}{2} z^T \bar{Q} z - c^T z.$$

Example 3. Finally, we mention the problem that motivated this note. It was described by Mangasarian and Musicant [6], who considered a QP formulation of the Huber regression problem. Given a matrix $A \in \mathsf{R}^{\ell \times d}$ and a vector $b \in \mathsf{R}^\ell$, we seek the vector $z \in \mathsf{R}^d$ that minimizes the objective function

$$\sum_{i=1}^{\ell} \rho((Az - b)_i), \tag{52}$$

where

$$\rho(t) = \begin{cases} \tfrac{1}{2} t^2, & |t| \leq \gamma, \\ \gamma |t| - \tfrac{1}{2} \gamma^2, & |t| > \gamma, \end{cases}$$

where γ is a positive parameter. By setting the derivative of (52) to zero, We can formulate this problem as an LCP by introducing variables $w, \lambda_1, \lambda_2 \in \mathsf{R}^\ell$ and writing

$$w - Az + b + \lambda^2 - \lambda^1 = 0, \tag{53a}$$
$$A^T w = 0, \tag{53b}$$
$$w + \gamma e \geq 0 \perp \lambda^1 \geq 0, \tag{53c}$$
$$-w + \gamma e \geq 0 \perp \lambda^2 \geq 0. \tag{53d}$$

We can write the problem as

$$\begin{bmatrix} -b \\ 0 \\ -\gamma e \\ -\gamma e \end{bmatrix} - \begin{bmatrix} I & -A & -I & I \\ A^T & 0 & 0 & 0 \\ I & 0 & 0 & 0 \\ -I & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} w \\ z \\ \lambda^1 \\ \lambda^2 \end{bmatrix} \in K^\circ, \tag{54a}$$

$$\begin{bmatrix} w \\ z \\ \lambda^1 \\ \lambda^2 \end{bmatrix} \in K, \tag{54b}$$

$$\left\langle \begin{bmatrix} w \\ z \\ \lambda^1 \\ \lambda^2 \end{bmatrix}, \begin{bmatrix} I & -A & -I & I \\ A^T & 0 & 0 & 0 \\ I & 0 & 0 & 0 \\ -I & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} w \\ z \\ \lambda^1 \\ \lambda^2 \end{bmatrix} - \begin{bmatrix} -b \\ 0 \\ -\gamma e \\ -\gamma e \end{bmatrix} \right\rangle = 0, \tag{54c}$$

where

$$K = \mathsf{R}^{\ell+d} \times \mathsf{R}^{2\ell}_{+}, \qquad K^\circ = \{0\} \times \mathsf{R}^{2\ell}_{-}. \tag{55}$$

Thus by defining

$$T = \mathsf{R}^{\ell} \times \{0\}, \;\; T^\perp = \{0\} \times \mathsf{R}^{2\ell+d}, \tag{56}$$

it is easy to verify that (17) and (18) are satisfied and that the properties (13) hold, with $Q = 0$ and $S = P_T$. Therefore the second statement of Theorem 3 holds, and (32) is

$$\begin{aligned} &\min \tfrac{1}{2} w^T w + \gamma e^T(\lambda_1 + \lambda_2), \\ &\text{subject to } \; w - Az + b + \lambda_2 - \lambda_1 = 0, \;\; \lambda_1 \geq 0, \;\; \lambda_2 \geq 0, \end{aligned} \tag{57a}$$

which is equivalent to the form given in [6, formula (9)]. Note that the naive QP formulation (7) would have many more constraints than this form.

Theorem 2 suggests another QP formulation for (54), one proposed by Li and Swetits [5]. From the form (22), we obtain

$$\begin{aligned} &\min \tfrac{1}{2} w^T w + b^T w, \\ &\text{subject to } \; -A^T w = 0, \;\; -\gamma e - w \leq 0, \;\; -\gamma e + w \leq 0, \end{aligned}$$

that is,

$$\min \tfrac{1}{2} w^T w + b^T w, \quad \text{subject to } \; -A^T w = 0, \;\; -\gamma e \leq w \leq \gamma e. \tag{58}$$

The second statement in Theorem 2 does not apply in this case, but we can still conclude that the primal-dual solution of (58) yields a solution of (54). In particular, the Lagrange multiplier vector for the constraint $-A^T w = 0$ yields a solution of (52).

Acknowledgments

I thank Olvi Mangasarian and Steve Robinson for their comments on the manuscript, and Todd Munson for pointers to relevant literature. I also thank Seetharam Gowda for pointing out an error in the original verison of Theorem 1 and for further pointers to the literature.

References

1. R. W. Cottle, J.-S. Pang, and R. E. Stone. *The Linear Complementarity Problem.* Academic Press, San Diego, 1992.
2. M. S. Gowda. Minimising quadratic functions over closed convex cones. *Bulletin of the Australian Mathematical Society*, 39:15–20, 1989.
3. M. S. Gowda and T. I. Seidman. Generalized linear complementarity problems. *Mathematical Programming, Series A*, 46:329–340, 1990.
4. Dong-Hui Li and Liqun Qi. A stabilized SQP method via linear equations. Technical report, Mathematics Department, University of New South Wales, 2000.
5. W. Li and J. J. Swetits. The linear ℓ_1 estimator and the Huber M-estimator. *SIAM Journal on Optimization*, 8:457–475, 1998.
6. O. L. Mangasarian and D. R. Musicant. Robust linear and support vector machines. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(9):1–6, 2000.
7. S. M. Robinson. A reduction method for variational inequalities. *Mathematical Programming, Series A*, 80:161–169, 1998.
8. S. M. Robinson. Composition duality and maximal monotonicity. *Mathematical Programming, Series A*, 85:1–13, 1999.
9. R. T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, N.J., 1970.
10. R. T. Rockafellar. Linear-quadratic programming and optimal control. *SIAM Journal on Control and Optimization*, 25:781–814, 1987.
11. R. T. Rockafellar. Computational schemes for large-scale problems in extended linear-quadratic programming. *Mathematical Programming*, 48:447–474, 1990.
12. S. J. Wright. Superlinear convergence of a stabilized SQP method to a degenerate solution. *Computational Optimization and Applications*, 11:253–275, 1998.