EXACT PENALTY FUNCTIONS FOR MATHEMATICAL PROGRAMS WITH LINEAR COMPLEMENTARITY CONSTRAINTS*

O.L. MANGASARIAN† AND J.S. PANG‡

Abstract. We establish a new general exact penalty function result for a constrained optimization problem and apply this result to a mathematical program with linear complementarity constraints.

Key words. Mathematical programs with complementarity constraints, exact penalty function, optimality conditions

AMS subject classifications. 90C30, 90C33

1. Introduction. A mathematical program with equilibrium constraints (MPEC) is a constrained optimization problem in which the essential constraints are defined by a variational inequality or complementarity system parametrized by a design variable. There has been a growing literature on this important mathematical programming problem which is central to many engineering design, economic equilibrium, multi-level game-theoretic and machine learning problems. The monograph [4] gives a comprehensive study of the MPEC and presents an extensive theory of exact penalty formulations and first- and second-order optimality conditions for this problem; some iterative algorithms for computing stationary points are also described therein.

In this paper, we establish some improved, exact penalty results for certain mathematical programs with linear complementarity constraints. Extension of the results to the case where these constraints contain free variables and associated equations (i.e. the case of mathematical programs with mixed linear complementarity constraints) is straightforward and not treated herein for the sake of simplicity. The new exact penalty results generalize previous ones for bilevel linear programs obtained in [1, 2] and complement those for mathematical programs with affine equilibrium constraints (MPAECs) in [5, 4]. The cornerstone of our new results is an improved *exact* penalty result for a constrained optimization problem that attains its minima at a certain *finite* set of points.

Specifically, we consider the following mathematical program with linear complementarity constraints:

$$\begin{array}{lll} & \text{minimize} & f(x,y) \\ (1) & \text{subject to} & (x,y) \in W \subseteq \Re^{n+m} \\ & & 0 \le y \perp w \equiv q + Nx + My \ge 0, \end{array}$$

* The work of the first author was based on research supported by National Science Foundation grant CCR-9322479. The work of the second author was based on research supported by National Science Foundation grant CCR-92137389.

† Computer Sciences Department, University of Wisconsin, Madison, Wisconsin 53706, U.S.A. Email: olvi@cs.wisc.edu

‡ Department of Mathematical Sciences, The Johns Hopkins University, Baltimore, Maryland 21218-2689, U.S.A. Email: jsp@vicp1.mts.jhu.edu

where $f : \Re^{n+m} \rightarrow \Re$ is a given real-valued function, W is a polyhedral set containing joint constraints on the variables (x, y), $q \in \Re^m$ is a given vector, and $N \in \Re^{m \times n}$ and $M \in \Re^{m \times m}$ are given matrices, and the symbol $\perp$ denotes orthogonality. This problem is a special case of the so-called MPAEC and includes as special cases traditional bilevel linear programs where x denotes the first-level variables and y denotes the variables (including slacks possibly) in a second-level linear-quadratic program. Its extension with free variables z and associated equations defined by (x, y, z) corresponds to the general MPAEC. Recently, several important problems in machine learning have been formulated as the problem (1) with a linear objective function f [3, 6, 7]; the acronym LPEC (for Linear Program with Equilibrium Constraints) was coined for this subclass of (1).

Our goal is to show that for certain classes of objective functions f, it is possible to obtain, by exact penalization of the orthogonality condition $y \perp w$, an equivalent formulation of the problem (1) as a linearly constrained nonlinear program.

2. A General Exact Penalty Result. In this section, we present an exact penalty function result for the following mathematical program:

$$\begin{array}{ll} \text{minimize} & \psi(u) \\ \text{subject to} & u \in S_1 \cap S_2, \end{array} \tag{2}$$

where $\psi : \Re^n \rightarrow \Re$ is a given real-valued function and S_1 and S_2 are two closed subsets of $\Re^n$. Typically S_1 is an "easy" set such as a polyhedral set, while S_2 is a "hard" set, possibly nonconvex. Let $\beta : \Re^n \rightarrow \Re_+$ be an arbitrary "exact penalty function" of the set S_2; that is,

$$\beta(u) \begin{cases} = 0 & \text{if } u \in S_2 \\ > 0 & \text{if } u \notin S_2. \end{cases}$$

For an arbitrary scalar $\alpha > 0$, define the penalty problem:

$$\begin{array}{ll} \text{minimize} & \psi(u) + \alpha\beta(u) \\ \text{subject to} & u \in S_1. \end{array} \tag{3}$$

Let $S^{\text{opt}}(\alpha)$ and S^{opt} denote, respectively, the set of global minimizers of the problem (3) and (2). In the following result, we postulate the existence of a finite subset of S_1 where the problem (3) attains its global minimum for all $\alpha > 0$ sufficiently large. In the application of this result to (1), S_1 is polyhedral with no lines and the finite subset is the set of extreme points of S_1.

THEOREM 2.1. *Suppose that there exists a finite subset S_0 of S_1 such that for all $\alpha > 0$,*

$$S_0(\alpha) \equiv S_0 \cap S^{\text{opt}}(\alpha) \neq \emptyset.$$

If $S_1 \cap S_2 \neq \emptyset$ then $S^{\text{opt}} \neq \emptyset$ and there exists a scalar $\bar{\alpha} > 0$ such that for all $\alpha \geq \bar{\alpha}$,

$$S_0(\alpha) \subseteq S^{\text{opt}} \subseteq S^{\text{opt}}(\alpha);$$

consequently,

$$S_0(\alpha) \;=\; S_0 \,\cap\, S^{\rm opt}.$$

Proof. We claim that there exists a scalar $\bar{\alpha} > 0$ such that for all $\alpha \geq \bar{\alpha}$, $S_0(\alpha) \subseteq S^{\rm opt}$. Clearly, for any $\alpha > 0$, if u is an element of $S_0(\alpha) \subseteq S_1$ that also belongs to S_2, then $u \in S^{\rm opt}$, because for $v \in S_1 \cap S_2$

$$\psi(v) \;=\; \psi(v) + \alpha\,\beta(v) \;\geq\; \psi(u) + \alpha\,\beta(u) \;=\; \psi(u).$$

Hence it suffices to show that for all α sufficiently large, every vector u in $S_0(\alpha)$ must belong to S_2. Assume for contradiction that this is false. Then for some sequence $\{\alpha_k\} \to \infty$, there exists $u^k \in S_0(\alpha_k)$ for each k such that $\beta(u^k) > 0$. Since each $S_0(\alpha_k)$ is a subset of the finite set S_0, we have

$$\beta_{\min} \;\equiv\; \inf_k \beta(u^k) \;>\; 0, \quad \text{and} \quad \rho \;\equiv\; \inf_k \psi(u^k) \;>\; -\infty.$$

Let $\bar{u} \in S_1 \cap S_2$ be arbitrary. We have for all k,

$$\psi(\bar{u}) \;\geq\; \psi(u^k) + \alpha_k\,\beta(u^k) \;\geq\; \rho + \alpha_k\,\beta_{\min}.$$

The right-hand term tends to ∞ as $k \to \infty$. This is a contradiction. Thus $S_0(\alpha) \subseteq S^{\rm opt}$ and $S^{\rm opt} \neq \emptyset$.

Next we show that for $\alpha \geq \bar{\alpha}$, $S^{\rm opt} \subseteq S^{\rm opt}(\alpha)$. For any such α, take $u^\alpha \in S_0(\alpha)$; let $\bar{u} \in S^{\rm opt}$ be arbitrary. Also let $u \in S_1$ be arbitrary. Since $u^\alpha \in S_1 \cap S_2$, we have, by the definition of u^α and $\bar{u}$,

$$\psi(u) + \alpha\,\beta(u) \;\geq\; \psi(u^\alpha) + \alpha\,\beta(u^\alpha) \;=\; \psi(u^\alpha) \;=\; \psi(\bar{u}) \;=\; \psi(\bar{u}) + \alpha\,\beta(\bar{u}).$$

Thus $\bar{u} \in S^{\rm opt}(\alpha)$.

To complete the proof, observe that

$$S_0 \,\cap\, S^{\rm opt} \;\subseteq\; S_0 \,\cap\, S^{\rm opt}(\alpha) \;=\; S_0(\alpha) \;\subseteq S_0 \,\cap\, S^{\rm opt}.$$

Thus equality holds throughout. □

The main difference between the above exact penalty function result and a traditional result of this type is that Theorem 2.1 does not assert that $S^{\rm opt}$ coincides with $S^{\rm opt}(\alpha)$ for all α sufficiently large. Instead, the theorem establishes the equality between these optimal sets intersected with the crucial finite set S_0. Needless to say, the finiteness of the set S_0 is essential to the proof (and the result).

3. Exact Penalty Results for the MPAEC. Problem (1) can be written in the following compact form:

$$\begin{array}{ll} \text{minimize} & f(x,y) \\ \text{subject to} & (x,y,w) \in V \subseteq \Re^{n+m+m} \\ & 0 \le y \perp w \ge 0, \end{array}$$

where

$$V \equiv \{(x,y,w) \in W \times \Re^m : (y,w) \ge 0, \; w = q + Nx + My\}$$

is a convex polyhedron containing the feasible region of (1). Assuming that the objective function f is concave and bounded below on the set V, we obtain the following exact penalty function result.

PROPOSITION 3.1. *Let W be a polyhedron in $\Re^{n+m}$ containing no lines and let $f : \Re^{n+m} \to \Re$ be a concave function. Assume that $V \neq \emptyset$ and f is bounded below on V. The following two statements are valid.*

(a) Problem (1) attains a finite minimum value, say $f_{\min}$.

(b) There exists a scalar $\bar{\alpha} > 0$ such that for all $\alpha \ge \bar{\alpha}$,

$$f_{\min} = \min\{\theta_\alpha(x,y,w) : (x,y,w) \in V\},$$

where

$$\theta_\alpha(x,y,w) \equiv f(x,y) + \alpha \sum_{i=1}^{m} \min(y_i, w_i);$$

moreover, every global minimizer of (1) is a global minimizer of θ_α on V; conversely, every global minimizer of θ_α on V that is an extreme point of V is a global minimizer of (1).

Proof. Clearly, V is a polyhedral convex set containing no lines; moreover for all $\alpha > 0$, the function θ_α is clearly concave. Hence by Corollary 32.3.4 in [9], for each $\alpha > 0$, the function θ_α attains its finite minimum at one of the finitely many extreme points of V. Moreover, since the feasible region of the problem (1) is equal to the union of finitely many convex polyhedra, none of which contain lines, it follows easily from the same quoted corollary that the minimum objective value of (1) is finite and attained. This establishes statement (a). By letting S_1 be V, S_0 be the finite set of extreme points of S_1,

$$S_2 \equiv \{(x,y,w) \in \Re^{n+m+m} : \min(y,w) = 0\},$$

and

$$\beta(x,y,w) \equiv \sum_{i=1}^{m} \min(y_i, w_i),$$

Theorem 2.1 readily yields assertion (b). □

The function θ_α is not differentiable because of the nondifferentiability of the min function. Noticing the elementary fact that for any two scalars a and b,

$$\min(a, b) = \min_{(\rho_1,\rho_2)\geq 0,\, \rho_1+\rho_2=1} (\rho_1 a + \rho_2 b),$$

we easily obtain a differentiable exact penalty formulation for (1). The following result is an immediate corollary Proposition 3.1 and does not require proof.

PROPOSITION 3.2. *Let W be a polyhedron in $\Re^{n+m}$ containing no lines and $f : \Re^{n+m} \rightarrow \Re$ be a concave function. Assume that $V \neq \emptyset$ and f is bounded below on V. Then there exists a scalar $\bar{\alpha} > 0$ such that for all $\alpha \geq \bar{\alpha}$,*

$$f_{\min} = \min \{ \phi_\alpha(x, y, w, r, s) : (x, y, w, r, s) \in V \times T \}$$

where

$$\phi_\alpha(x, y, w, r, s) \equiv f(x, y) + \alpha \sum_{i=1}^{m} (r_i y_i + s_i w_i)$$

and

$$T \equiv \{ (r, s) \in \Re^{2m}_{+} : r + s = e \},$$

with e is a vector of ones. Moreover, every global minimizer of (1) is a global minimizer of ϕ_α on $V \times T$; conversely, every global minimizer of ϕ_α on $V \times T$ that is an extreme point of $V \times T$ is a global minimizer of (1).

We now compare the above exact penalty results with those in Subsection 2.4.1 in [4] for the same problem (1). The essential difference lies in the assumptions made. Here, we assume that the objective function f is concave but we do not assume anything about the feasible region. In [4], there is no assumption of f but the feasible region is compact. We refer the reader to the cited monograph which discusses how the exact penalty results therein generalize those existing in the literature, in particular, those in [1, 2]. The papers [3, 6, 7] discuss the application of the class of linear programs with linear complementarity constraints to several important problems in machine learning. The exact penalty function formulation in Proposition 3.2 is obtained in the latter references. In [8, 3], successive linearization methods have been employed for solving these applied problems and found to be among the most effective methods for minimizing the penalty function ϕ_α on $V \times T$.

REFERENCES

[1] G. Anandalingam and D.J. White, "A solution method for the linear static Stackelberg problem using penalty function," *IEEE Transactions on Automatic Control* 35 (1990) 1170–1173.

[2] Z. Bi, P. Calamai, and A. Conn, "An exact penalty function approach for the linear bilevel programming problem," Technical Report #167-0-310789, Department of Systems Design and Engineering, University of Waterloo, Waterloo (1989).

[3] P.S. Bradley, O.L. Mangasarian, and W.N. Street, "Feature selection via mathematical programming", Mathematical Programming Technical Report 95-21, Department of Computer Science, University of Wisconsin, Madison (December 1995).

[4] Z.Q. Luo, J.S Pang and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press (1996).

[5] Z.Q. Luo, J.S Pang D. Ralph, and S.Q. Wu, "Exact penalization and stationarity conditions for mathematical programs with equilibrium constraints", *Mathematical Programming*, forthcoming.

[6] O.L. Mangasarian, "Misclassification minimization", *Journal of Global Optimization* 5 (1994) 309–323.

[7] O.L. Mangasarian, "Mathematical programming in machine learning", in "Nonlinear Optimization and Applications", G. Di Pillo and F. Giannessi, editors, Plenum Publishing , New York (1996), pp. 283-295.

[8] O.L. Mangasarian, "Machine learning via polyhedral concave minimization", in "Applied Mathematics and Parallel Computing - Festschrift for Klaus Ritter", H. Fischer and B. Riedmueller and S. Schaeffler, editors, Physica-Verlag A Springer-Verlag Company, Heidelberg (1996), pp. 175–188.

[9] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton (1970).