A Unified Augmented Lagrangian Approach to Duality and Exact Penalization

Article *in* Mathematics of Operations Research · August 2003

DOI: 10.1287/moor.28.3.533.16395 · Source: DBLP

Xiaoqi Yang
The Hong Kong Polytechnic University

MATHEMATICS OF OPERATIONS RESEARCH
Vol. 28, No. 3, August 2003, pp. 533–552
Printed in U.S.A.

A UNIFIED AUGMENTED LAGRANGIAN APPROACH TO DUALITY AND EXACT PENALIZATION

X. X. HUANG AND X. Q. YANG

In this paper, the existence of an optimal path and its convergence to the optimal set of a primal problem of minimizing an extended real-valued function are established via a generalized augmented Lagrangian and corresponding generalized augmented Lagrangian problems, in which no convexity is imposed on the augmenting function. These results further imply a zero duality gap property between the primal problem and the generalized augmented Lagrangian dual problem. A necessary and sufficient condition for the exact penalty representation in the framework of a generalized augmented Lagrangian is obtained. In the context of constrained programs, we show that generalized augmented Lagrangians present a unified approach to several classes of exact penalization results. Some equivalences among exact penalization results are obtained.

1. Introduction. It is well known that zero duality gap and exact penalty representation are important issues in mathematical programming. Duality theory and exact penalization for convex programming via ordinary Lagrangian have been well established. However, for nonconvex optimization problems, a nonzero duality gap may exist when using ordinary Lagrangian. In order to overcome this drawback, augmented Lagrangians were introduced in, e.g., Bestsekas (1982), Ioffe (1979), Rockafellar (1974, 1993) for constrained optimization problems. Recently in Rockafellar and Wets (1998) a convex augmenting function and a corresponding augmented Lagrangian were introduced for a primal problem of minimizing an extended real-valued function and, under mild conditions, a zero duality gap property and a necessary and sufficient condition for the exact penalization were established.

In this paper, we introduce a generalized augmented Lagrangian with a generalized augmenting function, which includes convex augmenting functions as a special case. We observe that by using a generalized augmenting function, "lower order" nonsmooth and nonconvex penalty functions used in, e.g., Luo and Pang (2000), Pang (1997), Luo et al. (1996), can be derived under the scheme of generalized augmented Lagrangian. For the existence of an exact penalty parameter, these nonsmooth and nonconvex penalty functions require only generalized analytic properties of functions involved, which are weaker than a regularity condition. These penalty functions have been intensively studied through so-called error bounds and successfully applied to the study of mathematical programs with equilibrium constraints; see, e.g., Luo and Pang (2000), Pang (1997), Luo et al. (1996). Furthermore, by transforming a constrained program into an equivalent constrained program, a class of nonlinear penalty functions studied in Rubinov et al. (1999a), Yang and Huang (2001) can also be derived using a generalized augmenting function. This class of nonlinear penalty functions admits a smaller least exact penalty parameter than the l_1 penalty function; see Rubinov et al. (1999a).

Received April 18, 2001; revised September 30, 2002.
MSC2000 subject classification. Primary: 90C30, 90C46.
OR/MS subject classification. Primary: Programming/nonlinear.
Key words. Generalized augmented Lagrangian, constrained program, duality, exact penalty function, nonlinear Lagrangian.

0364-765X/03/2803/0533/$05.00
1526-5471 electronic ISSN,

So the introduction of the generalized augmented Lagrangian allows one to make a unified study on the augmented Lagrangian in Rockafellar and Wets (1998), the nonsmooth and nonconvex penalty functions in Luo and Pang (2000), Pang (1997), Luo et al. (1996) and the nonlinear penalty functions in Rubinov et al. (1999a), Yang and Huang (2001). As also demonstrated by a simple example (see Example 4.1), for some primal problems, any augmented Lagrangian of Rockafellar amd Wets (1998) (i.e., a convex augmenting function being finite near the origin) cannot admit an exact penalty representation. However, the generalized augmented Lagrangians can admit exact penalty representations. In addition, our recent numerical tests on some generalized augmented Lagrangians for constrained programs with only equality constraints in Huang and Yang (to appear) show that usually a much smaller penalty parameter is needed to obtain the same or a better approximate optimal solution than the classical proximal augmented Lagrangian.

The existence and convergence of a path of optimal solutions generated by penalty type problems toward the optimal set is important for numerical methods Auslender (1999), Auslender et al. (1997), Yang and Huang (2001). In Rockafellar and Wets (1998), a zero duality gap result was established between the primal problem and its augmented Lagrangian dual problem. In this paper, under the same assumptions as in In Rockafellar and Wets (1998), we prove a deeper result in the framework of generalized augmented Lagrangian: the existence of a path of optimal solutions generated by generalized augmented Lagrangian problems and its convergence toward the optimal set of the primal problem. This result further implies the zero duality gap property between the generalized augmented Lagrangian dual problem and the primal problem.

A necessary and sufficient condition for the classical l_1 exact penalization was established in Burke (1991a, b). This result was recently extended for a class of nonlinear penalty functions (Rubinov et al. 1999a). A necessary and sufficient condition for exact penalty representation in the framework of augmented Lagrangian was given in Rockafellar and Wets (1998). In this paper, we obtain a necessary and sufficient condition for exact penalty representation in the framework of generalized augmented Lagrangian.

We apply the results obtained for the primal problem and its generalized augmented Lagrangian to a constrained program. In particular, we generalize an equivalence in terms of zero duality gap properties between the augmented Lagrangian dual scheme in Rockafellar and Wets (1998) and the nonlinear Lagrangian dual scheme (Rubinov et al. 1999a, b) established in Yang and Huang (2001) under the condition that the generalized augmenting function is assumed to be continuous at the origin, which is weaker than the local Lipschitzianness requirement on the augmenting function in Yang and Huang (2001). Various penalty functions may be used to solve a constrained program. Some may be exact while others may not. It is interesting to compare different penalty functions in the sense that the exactness of one penalty function may imply that of the other. In this paper, we investigate the relationships of exact penalty representations among generalized augmented Lagrangian functions, nonsmooth and nonconvex penalty functions in Luo and Pang (2000), Pang (1997), Luo et al. (1996) and nonlinear penalty functions in Rubinov et al. (1999a).

The outline of this paper is as follows. In §2, we introduce a generalized augmented Lagrangian and obtain several penalty functions for an inequality and equality constrained program as special cases. We also establish the existence of a path of optimal solutions and its convergence to the optimal set of the primal problem and consequently derive a zero duality gap property for the primal problem and its generalized augmented Lagrangian dual problem. In §3, we obtain necessary and sufficient conditions for an exact penalty representation in the framework of generalized augmented Lagrangians. Finally, in §4, we apply the results obtained in §§2 and 3 to a constrained program and its generalized augmented Lagrangian problems.

2. Generalized augmented Lagrangian and zero duality gap. Let $\bar{R} = R \cup \{+\infty, -\infty\}$ and φ: $R^n \to \bar{R}$ be an extended real-valued function. Consider the primal problem

$$\inf_{x \in R^n} \varphi(x). \tag{1}$$

A function $\bar{f}$: $R^n \times R^m \to \bar{R}$ is said to be a *dualizing parameterization function* for φ if $\varphi(x) = \bar{f}(x, 0)$, $\forall\, x \in R^n$.

DEFINITION 2.1 (ROCKAFELLAR AND WETS 1998). (i) Let $X \subset R^n$ be a closed subset and f: $X \to \bar{R}$ be an extended real-valued function. The function f is said to be *level-bounded* on X if, for any $\alpha \in R$, the set $\{x \in X\colon f(x) \le \alpha\}$ is bounded.

(ii) A function $\bar{f}$: $R^n \times R^m \to \bar{R}$ with value $\bar{f}(x, u)$ is said to be *level-bounded in x locally uniform in u* if, for each $\bar{u} \in R^m$ and $\alpha \in R$, there exists a neighborhood $U(\bar{u})$ of $\bar{u}$ along with a bounded set $D \subset R^n$, such that $\{x \in R^n\colon \bar{f}(x, u) \le \alpha\} \subset D$ for any $u \in U(\bar{u})$.

DEFINITION 2.2. A function σ: $R^m \to R_+ \cup \{+\infty\}$ is said to be a *generalized augmenting function* if it is proper, lower semicontinuous (lsc, for short), level-bounded on R^m, $\arg\min_y \sigma(y) = \{0\}$, and $\sigma(0) = 0$.

It is worth noting that this definition of a generalized augmenting function is different from that of the augmenting function given in Definition 11.55 of Rockafellar and Wets (1998) in that no convexity requirement is imposed. By Definition 2.2, any augmenting function according to Definition 11.55 of Rockafellar and Wets (1998) is a generalized augmenting function.

Let $u = (u_1, \ldots, u_m) \in R^m$ and

$$\|u\|_p = \begin{cases} \left(\displaystyle\sum_{j=1}^{m} |u_j|^p\right)^{1/p}, & \text{if } 1 \le p < +\infty, \\ \max\{|u_1|, \ldots, |u_m|\}, & \text{if } p = +\infty. \end{cases}$$

If σ is an augmenting function in the sense of Rockafellar and Wets (1998) then, for any $\gamma > 0$, σ^γ is a generalized augmenting function. In particular,

(a) if $\sigma(u) = \|u\|_1$, then for any $\gamma > 0$, $\sigma^\gamma(u) = \|u\|_1^\gamma$ is a generalized augmenting function;

(b) if $\sigma(u) = \|u\|_\infty$, then $\sigma^\gamma(u) = \|u\|_\infty^\gamma$ is a generalized augmenting function for any $\gamma > 0$.

(c) if $\sigma(u) = \sum_{j=1}^m |u_j|^\gamma$, where $\gamma > 0$, then $\sigma(u)$ is a generalized augmenting function.

It is clear that none of these three classes (a), (b), and (c) of generalized augmenting functions is convex when $\gamma \in (0, 1)$; namely, none of them is an augmenting function.

DEFINITION 2.3. Consider the primal problem (1). Let $\bar{f}$ be a dualizing parameterization function for φ, and let σ be a generalized augmenting function. The generalized augmented Lagrangian (with parameter $r > 0$) l: $R^n \times R^m \times (0, +\infty) \to \bar{R}$ is defined by

$$l(x, y, r) = \inf\big\{\bar{f}(x, u) - \langle y, u\rangle + r\sigma(u)\colon u \in R^m\big\}, \qquad x \in R^n, y \in R^m, r > 0,$$

where $\langle y, u\rangle$ denotes the inner product.

EXAMPLE 2.1. Consider the following constrained program

$$\text{(P)} \qquad \begin{aligned} \inf \quad & f(x) \\ \text{s.t.} \quad & x \in X, \\ & g_j(x) \le 0, \qquad j = 1, \ldots, m_1, \\ & g_j(x) = 0, \qquad j = m_1 + 1, \ldots, m, \end{aligned}$$

where $X \subset R^n$ is a nonempty and closed set, f, g_j: $X \to R^1$, $j = 1, \ldots, m_1$ are lsc and g_j: $X \to R^1$, $j = m_1 + 1, \ldots, m$ are continuous. We assume without loss of generality that $\inf_{x\in X} f(x) > 0$. Otherwise, we may replace $f(x)$ by $\exp(f(x)) + 1$ and the resulting constrained program has the same set of (locally) optimal solutions as (P). Denote by X_0 the set of feasible solutions of (P), i.e.,

$$X_0 = \left\{x \in X\colon g_j(x) \le 0, j = 1, \ldots, m_1, g_j(x) = 0, j = m_1 + 1, \ldots, m\right\},$$

and by M_P the optimal value of problem (P). Denote $g^+(x) = \max\{g(x), 0\}$.

Let

$$\varphi(x) = \begin{cases} f(x), & \text{if } x \in X_0, \\ +\infty, & \text{if } x \in R^n \backslash X_0. \end{cases} \tag{2}$$

It is clear that (P) is equivalent to the following problem (P') in the sense that the two problems have the same set of (locally) optimal solutions and the same optimal value:

$$\inf_{x\in R^n} \varphi(x). \tag{P$'$}$$

Define the dualizing parameterization function:

$$\bar{f}_P(x, u) = f(x) + \delta_{R_-^{m_1} \times \{0_{m-m_1}\}}(G(x) + u) + \delta_X(x), \qquad x \in R^n, u \in R^m, \tag{3}$$

where 0_{m-m_1} is the origin of R^{m-m_1}, $G(x) = (g_1(x), \ldots, g_m(x))$, and δ_D is the indicator function of the set D, i.e.,

$$\delta_D(z) = \begin{cases} 0, & \text{if } z \in D, \\ +\infty, & \text{otherwise.} \end{cases}$$

Thus, a class of generalized augmented Lagrangians for (P) with the dualizing parameterization function $\bar{f}_P$ defined by (3) can be expressed as

$$l_P(x, y, r) = \inf\left\{\bar{f}_P(x, u) - \langle y, u\rangle + r\sigma(u)\colon u \in R^m\right\}, \tag{4}$$

where σ is a generalized augmenting function.

When $\sigma(u) = \frac{1}{2}\|u\|_2^2$, it follows from Example 11.57 in Rockafellar and Wets (1998) (setting $D = R_-^{m_1} \times 0_{m-m_1}$) that

$$l_P(x, y, r) = \begin{cases} f(x) + (r/2)\left[\sum_{j=1}^{m_1} (r^{-1}y_j + g_j(x))^{+2} - \sum_{j=1}^{m_1} (r^{-1}y_j)^2\right] & \\ \quad + \sum_{j=m_1+1}^{m} \left\{y_j g_j(x) + (r/2)g_j^2(x)\right\}, & \text{if } x \in X, \\ +\infty, & \text{otherwise,} \end{cases}$$

which is just the augmented Lagrangian defined in formula (6.7) in Rockafellar (1993).

EXAMPLE 2.2. (i) For the general constrained program (P), taking $\sigma(u) = \|u\|_1^\gamma = (\sum_{j=1}^m |u_j|)^\gamma$, $u \in R^m$, one can derive from the generalized augmented Lagrangian $l_\gamma(x, y, r)$ (for (P)) defined by (4) that

$$l_\gamma(x, 0, r) = \begin{cases} \inf\{\bar{f}_P(x, u) + \delta_{R_-^{m_1} \times \{0_{m-m_1}\}}(G(x) + u) + r\|u\|_1^\gamma\colon u \in R^m\}, & \text{if } x \in X, \\ +\infty, & \text{otherwise,} \end{cases}$$

$$= \begin{cases} f(x) + r\left(\sum_{j=1}^{m_1} g_j^+(x) + \sum_{j=m_1+1}^{m} |g_j(x)|\right)^\gamma, & \text{if } x \in X, \\ +\infty, & \text{otherwise.} \end{cases}$$

This gives rise to the penalty function in Luo et al. (1996) for an ordinary constrained mathematical program. Specifically, the following penalty function was considered in Luo et al. (1996):

$$q_\gamma(x,r) = f(x) + r\left(\sum_{j=1}^{m_1} g_j^+(x) + \sum_{j=m_1+1}^{m} |g_j(x)|\right)^\gamma, \qquad x \in X, r > 0. \tag{5}$$

(ii) Choose $\sigma(u) = \|u\|_\infty^\gamma$, where $\gamma > 0$. Then it can be easily computed from the generalized augmented Lagrangian for (P) that

$$l_\infty(x,0,r) = \begin{cases} f(x) + r\left[\max\{g_1^+(x),\dots,g_{m_1}^+(x), |g_{m_1+1}(x)|,\dots,|g_m(x)|\}\right]^\gamma, & \text{if } x \in X,\\ +\infty, & \text{otherwise.}\end{cases}$$

This leads us to the penalty function in Pang (1997):

$$f(x) + r\left[\max\{g_1^+(x),\dots,g_{m_1}^+(x), |g_{m_1+1}|,\dots,|g_m(x)|\}\right]^\gamma, \qquad x \in X, r > 0.$$

It is worth mentioning that these two classes of penalty functions have been intensively studied in the literature through the study of the so-called error bounds (see, e.g., Luo et al. 1996, Luo and Pang 2000, Pang 1997). These two classes of penalty functions are called penalty functions of order γ. As noted in Luo et al. (1996), the classical l_1 exact penalty function (called exact penalization of order 1) requires a regularity condition while the two classes of exact penalty functions of order γ with $\gamma \in (0,1)$ require only some generalized analytic properties of the constraint set. These analytic properties are very broad and often can be satisfied by sets that fail the regularity assumption. In other words, the two classes of penalty functions of order γ with $\gamma \in (0,1)$ derived above require much weaker conditions than the classical l_1 penalty function to guarantee their exact penalization property. A necessary and sufficient condition for these two classes of exact penalty functions of order γ with $\gamma \in (0,1)$ will be derived in §4.2 below.

EXAMPLE 2.3. Let $\gamma > 0$. Consider the problem (P). Noticing the assumption that $\inf_{x\in X} f(x) > 0$, the constrained program (P) is equivalent to the following constrained program:

$$(P_\gamma) \qquad \begin{array}{ll} \inf & f^\gamma(x) \\ \text{s.t.} & x \in X, \\ & g_j(x) \le 0, \qquad j = 1,\dots,m_1, \\ & g_j(x) = 0, \qquad j = m_1+1,\dots,m. \end{array}$$

Moreover, the optimal value of problem (P_γ) is exactly M_P^γ.

Take $\sigma(u) = \sum_{j=1}^m |u_j|^\gamma$, $u = (u_1,\dots,u_m) \in R^m$. Applying the procedure of defining the generalized augmented Lagrangians for the constrained program (P) described above to (P_γ), we can define a class of generalized augmented Lagrangians for (P_γ):

$$l'_\gamma(x,y,r) = \inf\{\bar f_\gamma(x,u) - \langle y,u\rangle + r\sigma(u)\colon u \in R^m\},$$

where

$$\bar f_\gamma(x,u) = f^\gamma(x) + \delta_{R_-^{m_1}\times\{0_{m-m_1}\}}(G(x)+u) + \delta_X(x), \qquad x \in R^n, u \in R^m.$$

Analogously, we get

$$l'_\gamma(x,0,r) = \begin{cases} f^\gamma(x) + r\left(\sum_{j=1}^{m_1} g_j^{+\gamma}(x) + \sum_{j=m_1+1}^{m} |g_j(x)|^\gamma\right), & \text{if } x \in X,\\ +\infty, & \text{otherwise.}\end{cases}$$

It is worth noting that $l'_\gamma(x, 0, r)$ is closely related to a class of nonlinear Lagrangians for (P) (see Rubinov et al. 1999a, b). In particular, the following nonlinear penalty function for (P) was considered in Rubinov (1999a):

$$(6)\quad L_\gamma(x, r) = \left[f^\gamma(x) + r\left(\sum_{j=1}^{m_1} g_j^{+\gamma}(x) + \sum_{j=m_1+1}^{m} |g_j(x)|^\gamma\right)\right]^{1/\gamma}, \qquad x \in X,\, r \in (0, +\infty).$$

It was shown in Rubinov (1999a) that, compared to the l_1 penalty function, the penalty function $L_{1/2}(x, r)$ requires weaker conditions to guarantee its exact penalization property and admits a smaller least exact penalty parameter. These properties show that the $L_{1/2}$ penalty function has some advantages over the l_1 penalty function in both theory and computation.

Now we return to the discussion of the generalized augmented Lagrangian for the primal problem (1).

Let $l(x, y, r)$ be a generalized augmented Lagrangian for (1). The *generalized augmented Lagrangian dual function* corresponding to l is defined by

$$(7)\qquad \bar{\psi}(y, r) = \inf\{l(x, y, r)\colon x \in R^n\}, \qquad y \in R^m,\, r > 0.$$

The *generalized augmented Lagrangian dual problem* is defined as

$$(8)\qquad \sup_{(y,r)\in R^m\times(0,+\infty)} \bar{\psi}(y, r).$$

Define the perturbation function of (1) by

$$(9)\qquad p(u) = \inf\{\bar{f}(x, u)\colon x \in R^n\}.$$

Then $p(0)$ is just the optimal value of the problem (1).

The following proposition summarizes some basic properties of the generalized augmented Lagrangian, which will be useful in the sequel. Its proof is elementary and omitted.

PROPOSITION 2.1. *For any dualizing parameterization and any generalized augmenting function, we have*

(i) *the generalized augmented Lagrangian $l(x, y, r)$ is concave, upper semicontinuous in (y, r) and nondecreasing in r;*

(ii) *weak duality holds:*

$$(10)\qquad \bar{\psi}(y, r) \le p(0), \qquad \forall\, (y, r) \in R^m \times (0, +\infty).$$

In what follows, under mild conditions, we establish the existence of a path of optimal solutions generated by the generalized augmented Lagrangian problems and its convergence to the optimal solution set of the primal problem. As a result, the convergence itself implies a zero duality gap property between the primal problem (1) and its generalized augmented Lagrangian dual problem (8).

Consider the primal problem (1) and its generalized augmented Lagrangian problem:

$$P(y, r)\qquad \inf_{(x,u)\in R^n\times R^m} \left\{\bar{f}(x, u) + r\sigma(u) - \langle y, u\rangle\right\}.$$

Note that $P(y, r)$ is the same as the problem of evaluating the generalized augmented Lagrangian dual function $\bar{\psi}(y, r)$. Let S and $V(y, r)$ denote the optimal solution sets of the problems (1) and $P(y, r)$, respectively. Recall that $p(0)$ and $\bar{\psi}(y, r)$ are the optimal values of the problems (1) and $P(y, r)$, respectively.

THEOREM 2.1 (OPTIMAL PATH AND ZERO DUALITY GAP). *Consider the primal problem* (1), *its generalized augmented Lagrangian problem* $P(y, r)$ *and generalized augmented Lagrangian dual problem* (8). *Assume that* φ *is proper, and that its dualizing parameterization function* $\bar{f}(x, u)$ *is proper, lsc, and level-bounded in* x *locally uniform in* u. *Suppose that there exists* $(\bar{y}, \bar{r}) \in R^m \times (0, +\infty)$ *such that*

$$\inf\{l(x, \bar{y}, \bar{r}): x \in R^n\} > -\infty.$$

Then

(i) S *is nonempty and compact*;

(ii) *for any* $r \geq \bar{r}+1$, $V(\bar{y}, r)$ *is nonempty and compact, where* $(\bar{y}, \bar{r})$ *is a pair meeting*

$$\bar{f}(x, u)+\bar{r}\sigma(u)-\langle \bar{y}, u\rangle \geq m_0, \qquad \forall\, x \in R^n, u \in R^m, \tag{11}$$

for some $m_0 \in R$;

(iii) *for each selection* $(x(r), u(r)) \in V(\bar{y}, r)$ *with* $r \geq \bar{r}+1$, *the optimal path* $\{(x(r), u(r))\}$ *is bounded and its limit points take the form* $(x^*, 0)$, *where* $x^* \in S$;

(iv) $p(0) = \lim_{r\to+\infty} \bar{\psi}(\bar{y}, r)$;

(v) *zero duality gap holds*:

$$p(0) = \sup_{(y,\, r)\in R^m\times(0,+\infty)} \bar{\psi}(y, r).$$

PROOF. (i) From the assumption that the dualizing parametrization function $\bar{f}(x, u)$ is lsc, and level-bounded in x, we see that φ is lsc and level-bounded. It follows that S is nonempty and compact.

(ii) Let $\bar{x} \in R^n$ be such that $-\infty < \varphi(\bar{x}) < +\infty$. Let

$$U(r) = \big\{(x, u) \in R^n \times R^m: \bar{f}(x, u)+r\sigma(u)-\langle \bar{y}, u\rangle \leq \varphi(\bar{x})\big\}.$$

We prove that $U(\bar{r}+1)$ is a compact set. Suppose to the contrary that $\exists (x^k, u^k) \in U(\bar{r}+1)$ such that $\|(x^k, u^k)\| \to +\infty$. Since $(x^k, u^k) \in U(\bar{r}+1)$, we have

$$\bar{f}(x^k, u^k)+\bar{r}\sigma(u^k)-\langle \bar{y}, u^k\rangle+\sigma(u^k) \leq \varphi(\bar{x}). \tag{12}$$

This, combined with (11), yields

$$\sigma(u^k) \leq \varphi(\bar{x})-m_0.$$

By the level-boundedness of σ, we see that $\{u^k\}$ is bounded. Without loss of generality, suppose that $u^k \to \bar{u}$. From (12), we have

$$\bar{f}(x^k, u^k) \leq \varphi(\bar{x})+\langle \bar{y}, u^k\rangle \leq \varphi(\bar{x})+t_0, \tag{13}$$

for some $t_0 > 0$. As $\bar{f}(x, u)$ is level-bounded in x locally uniform in u, we deduce from (13) that $\{x^k\}$ is bounded. It follows that $\{(x^k, u^k)\}$ is bounded, a contradiction. Thus, $U(\bar{r}+1)$ is compact. Since $U(r) \subset U(\bar{r}+1)$ and $U(r)$ is closed, $U(r)$ is compact for $r \geq \bar{r}+1$.

In addition, it is clear that $U(r) \neq \varnothing, \forall\, r > 0$ since $(\bar{x}, 0) \in U(r)$. So $U(r)$ is nonempty and compact whenever $r \geq \bar{r}+1$. As a result, the problem $P(\bar{y}, r)$ has a solution whenever $r \geq \bar{r}+1$. Thus, the solution set $V(\bar{y}, r) \subset U(\bar{r}+1)$ is nonempty and compact for any $r \geq \bar{r}+1$.

(iii) Let $(x(r), u(r)) \in V(\bar{y}, r)$ with $r \geq \bar{r}+1$. Since $(x(r), u(r)) \in U(\bar{r}+1)$ and $U(\bar{r}+1)$ is compact, it follows that $\{(x(r), u(r)\}$ is bounded. Suppose that (x^*, u^*) is a limit point of $\{(x(r), u(r))\}$. Then, there exists $\bar{r}+1 < r_k \to +\infty$ and $(x(r_k), u(r_k)) \in V(\bar{y}, r_k)$ such that $(x(r_k), u(r_k)) \to (x^*, u^*)$.

Arbitrarily fix an $x \in R^n$. It is clear that

$$\bar{f}(x(r_k), u(r_k)) + r_k\sigma(u(r_k)) - \langle \bar{y}, u(r_k)\rangle \le \bar{f}(x, 0) = \varphi(x). \tag{14}$$

Inequality (14), together with (11), gives

$$(r_k - \bar{r})\sigma(u(r_k)) \le \varphi(x) - m_0, \qquad \forall k.$$

Thus,

$$\sigma(u(r_k)) \le \frac{\varphi(x) - m_0}{r_k - \bar{r}}.$$

This inequality, together with the lsc property of σ, gives

$$\sigma(u^*) \le \liminf_{k\to+\infty} \sigma(u(r_k)) = 0.$$

Therefore, $u^* = 0$.

Note that (14) implies

$$\bar{f}(x(r_k), u(r_k)) - \langle \bar{y}, u(r_k)\rangle \le \varphi(x). \tag{15}$$

Taking the lower limit in (15) as $k \to +\infty$ gives

$$\varphi(x^*) = f(x^*, 0) \le \liminf_{k\to+\infty} \bar{f}(x(r_k), u(r_k)) - \langle \bar{y}, u(r_k)\rangle \le \varphi(x).$$

By the arbitrariness of $x \in R^n$, we conclude that $x^* \in S$. So (iii) is proved.

(iv) We need only show that, for each sequence $r_k \to +\infty$, $\bar{\psi}(\bar{y}, r_k) \to p(0)$. From (iii), suppose that there exists $(x(r_k), u(r_k)) \in V(\bar{y}, r_k)$ such that

$$\bar{\psi}(\bar{y}, r_k) = \bar{f}(x(r_k), u(r_k)) - \langle \bar{y}, u(r_k)\rangle + r_k\sigma(u(r_k)),$$

and that, without loss of generality,

$$(x(r_k), u(r_k)) \to (x^*, 0) \in S \times \{0\}. \tag{16}$$

Consequently,

$$\bar{\psi}(\bar{y}, r_k) \ge \bar{f}(x(r_k), u(r_k)) - \langle \bar{y}, u(r_k)\rangle.$$

This, combined with the lsc of $\bar{f}(\cdot, \cdot)$ and (16), yields

$$\begin{aligned}\liminf_{k\to+\infty} \bar{\psi}(\bar{y}, r_k) &\ge \liminf_{k\to+\infty} \bar{f}(x(r_k), u(r_k)) - \langle \bar{y}, u(r_k)\rangle \\ &\ge \bar{f}(x^*, 0) = \varphi(x^*) = p(0). \end{aligned}\tag{17}$$

Note also, by Proposition 2.1(ii),

$$\limsup_{k\to+\infty} \bar{\psi}(\bar{y}, r_k) \le p(0). \tag{18}$$

(17) together with (18) imply:

$$\lim_{k\to+\infty} \bar{\psi}(\bar{y}, r_k) = p(0).$$

So (iv) has been proved.

(v) From (iv), we deduce

$$\sup_{(\bar{y}, r)\in R^m \times (0,+\infty)} \bar{\psi}(\bar{y}, r) \ge p(0). \tag{19}$$

This, together with Proposition 2.1(ii), yields (v). The proof is complete. $\square$

REMARK 2.1. The properness of φ and the lsc of $\bar{f}$ are necessary for the validity of Theorem 11.59 in Rockafellar and Wets (1998) as demonstrated by Examples 2.1 and 2.2 in Huang and Yang (2001). These two counterexamples can also be used to show that the properness of φ and the lsc of $\bar{f}$ are also essential to guarantee the validity of Theorem 2.1.

3. Exact penalty representation. In this section, we establish exact penalization representation results in the framework of generalized augmented Lagrangians, extending Theorem 11.61 in Rockafellar and Wets (1998).

DEFINITION 3.1 (EXACT PENALTY REPRESENTATION). Consider the problem (1). Let p be the perturbation function defined by (9). Let l be a generalized augmented Lagrangian for (1). A vector $\bar{y} \in R^m$ is said to support an exact penalty representation for the problem (1) if there exists $\bar{r} > 0$ such that

$$p(0) = \inf_{x \in R^n} l(x, \bar{y}, r), \qquad \forall\, r \geq \bar{r}, \tag{20}$$

and

$$\arg\min_x \varphi(x) = \arg\min_x \bar{l}(x, \bar{y}, r), \qquad \forall\, r \geq \bar{r}. \tag{21}$$

THEOREM 3.1. *In the framework of the generalized augmented Lagrangian l, the following statements are true*:

(i) *If $\bar{y}$ supports an exact penalty representation for the problem* (1), *then there exist $\bar{r} > 0$ and a neighborhood W of $0 \in R^m$ such that*

$$p(u) \geq p(0) + \langle \bar{y}, u \rangle - \bar{r}\sigma(u), \qquad \forall\, u \in W.$$

(ii) *The converse of* (i) *is true if*

(a) $p(0)$ *is finite*;

(b) *there exists $\bar{r}' > 0$ such that*

$$\inf\{\bar{f}(x, u) - \langle \bar{y}, u \rangle + \bar{r}'\sigma(u) \colon (x, u) \in R^n \times R^m\} > -\infty;$$

(c) *there exist $\tau > 0$ and $N > 0$ such that $\sigma(u) \geq \tau\|u\|$ when $\|u\| \geq N$.*

PROOF. Since $\bar{y}$ supports an exact penalty representation, there exists $\bar{r} > 0$ such that (20) holds with $r = \bar{r}$, i.e.,

$$p(0) = \inf\{l(x, \bar{y}, \bar{r}) \colon x \in R^n\} = \inf\{\bar{f}(x, u) - \langle \bar{y}, u \rangle + \bar{r}\sigma(u) \colon (x, u) \in R^n \times R^m\}.$$

Consequently,

$$p(0) \leq \bar{f}(x, u) - \langle \bar{y}, u \rangle + \bar{r}\sigma(u), \qquad \forall\, x \in R^n, u \in R^m,$$

implying

$$p(0) \leq p(u) - \langle \bar{y}, u \rangle + \bar{r}\sigma(u), \qquad \forall\, u \in R^m.$$

This proves (i).

It is evident from the proof of Theorem 11.61 in Rockafellar and Wets (1998) that (ii) is true. □

REMARK 3.1. In Rockafellar and Wets (1998), σ was assumed to be proper, lsc, convex and $\arg\min_y \sigma(y) = \{0\}$. As noted in Rockafellar and Wets (1998), σ is level-coercive. It follows that this assumption implies the existence of $\tau > 0$ and $N > 0$ satisfying $\sigma(u) \geq \tau\|u\|$ when $\|u\| \geq N$.

For the special case where $\bar{y} = 0$ supports an exact penalty representation for the problem (1), the following result shows that weaker conditions (without (ii)(c) in Theorem 3.1) are required.

THEOREM 3.2. *In the framework of the generalized augmented Lagrangian l, the following statements are true*:

(i) *If* $\bar{y} = 0$ *supports an exact penalty representation, then there exist* $\bar{r} > 0$ *and a neighborhood* W *of* $0 \in R^m$ *such that*

$$p(u) \geq p(0) - \bar{r}\sigma(u), \qquad \forall\, u \in W. \tag{22}$$

(ii) *The converse of* (i) *is true if*

(a) $p(0)$ *is finite*;

(b) *there exist* $\bar{r}' > 0$ *and* $m^* \in R$ *such that* $\bar{f}(x, u) + \bar{r}'\sigma(u) \geq m^*, \forall\, x \in R^n, u \in R^m$.

PROOF. (i) follows from Theorem 3.1(i). We need only to prove (ii). Assume that (22) holds.

First we prove (20) by contradiction. Suppose by the weak duality (10) that there exists $0 < r_k \to +\infty$ with

$$p(0) > \inf_{x \in R^n} l(x, 0, r_k).$$

Then there exist $x^k \in R^n$ and $u^k \in R^m$ such that

$$\begin{aligned} p(0) &> \bar{f}(x^k, u^k) + r_k\sigma(u^k) \\ &= \bar{f}(x^k, u^k) + \bar{r}'\sigma(u^k) + (r_k - \bar{r}')\sigma(u^k) \\ &\geq m^* + (r_k - \bar{r}')\sigma(u^k). \end{aligned} \tag{23}$$

The level-boundedness of σ implies that $\{u^k\}$ is bounded. Assume, without loss of generality, that $u^k \to \bar{u}$. It follows from (24) that

$$\sigma(\bar{u}) \leq \liminf_{k \to +\infty} \sigma(u^k) \leq \lim_{k \to +\infty} \frac{p(0) - m^*}{r_k - \bar{r}'} = 0.$$

Thus $\bar{u} = 0$. From the first inequality in (24), we deduce that

$$p(0) > p(u^k) + r_k\sigma(u^k), \qquad \forall\, k. \tag{24}$$

Since $u^k \to 0$, we conclude that (24) contradicts (22). As a result, there exists $\bar{r} > \max(\bar{r}', \hat{r})$ such that (20) holds.

For any $x^* \in \arg\min_x \varphi(x)$, we have from (22) that

$$\varphi(x^*) = l(x^*, 0, r) = \inf_{x \in R^n} l(x, 0, r), \qquad r \geq \bar{r}.$$

Therefore, $x^* \in \arg\min_x \bar{l}(x, 0, r)$. This shows that

$$\arg\min_x \varphi(x) \subseteq \arg\min_x \bar{l}(x, 0, r), \qquad r \geq \bar{r}.$$

Now we show that there exists $r^* > \bar{r} + 1 > 0$ such that

$$\arg\min_x \bar{l}(x, 0, r) \subseteq \arg\min_x \varphi(x), \qquad \forall\, r > r^*.$$

Suppose to the contrary that there exists $\bar{r} + 1 < r_k \uparrow +\infty$ and $x^k \in \arg\min_x \bar{l}(x, 0, r_k)$ such that $x^k \notin \arg\min_x \varphi(x), \forall\, k$. Then

$$\varphi(x^k) > p(0), \qquad \forall\, k. \tag{25}$$

For each fixed k, by the definition of $l(x^k, 0, r_k)$, $\exists\{u^{k,i}\} \subset R^m$ with

$$\bar{f}(x^k, u^{k,i}) + r_k\sigma(u^{k,i}) \to \bar{l}(x^k, 0, r_k) = p(0),$$

as $i \to +\infty$, namely,

$$\bar{f}(x^k, u^{k,i}) + \bar{r}'\sigma(u^{k,i}) + (r_k - \bar{r}')\sigma(u^{k,i}) \to p(0),$$

as $i \to +\infty$. It follows that $\{(r_k - \bar{r}')\sigma(u^{k,i})\}_{i=1}^{+\infty}$ is bounded since $\bar{f}(x^k, u^{k,i}) + \bar{r}'\sigma(u^{k,i}) \geq m^*$. As σ is level-bounded, we know that $\{u^{k,i}\}_{i=1}^{+\infty}$ is bounded. Without loss of generality, assume that $u^{k,i} \to \bar{u}^k$. Then

$$\bar{f}(x^k, \bar{u}^k) + r_k\sigma(\bar{u}^k) \leq \liminf_{i\to+\infty} \bar{f}(x^k, u^{k,i}) + r_k\sigma(u^{k,i}) = p(0). \tag{26}$$

Hence,

$$\bar{f}(x^k, \bar{u}^k) + \bar{r}'\sigma(\bar{u}^k) + (r_k - \bar{r}')\sigma(\bar{u}^k) \leq p(0).$$

So

$$(r_k - \bar{r}')\sigma(\bar{u}^k) \leq p(0) - m^*. \tag{27}$$

Again, by the level-boundedness of σ, we see that $\{\bar{u}^k\}$ is bounded. Suppose, without loss of generality, that $\bar{u}^k \to \bar{u}$. Then, from (27), we obtain

$$\sigma(\bar{u}) \leq \liminf_{k\to+\infty} \sigma(\bar{u}^k) \leq \lim_{k\to+\infty} \frac{p(0) - m^*}{r_k - \bar{r}'} = 0.$$

So we know that $\bar{u}^k \to 0$. Note from (25) that $\bar{u}^k \neq 0, \forall\, k$. As a result, (26) contradicts (22). The proof is complete. □

4. Applications to constrained programs. In this section, we apply the results established in §§2 and 3 for the primal problem (1) to the constrained optimization problem (P).

4.1. Zero duality gap property. Consider the constrained program (P) and its generalized augmented Lagrangian l_P given in §2. The corresponding generalized augmented Lagrangian dual function for (P) is

$$\bar{\psi}_P(y, r) = \inf\{l_P(x, y, r)\colon x \in R^n\}, \qquad y \in R^m, r \in (0, +\infty). \tag{28}$$

The corresponding *generalized augmented Lagrangian dual problem*is defined as

$$(D_A) \qquad \sup_{(y,r)\in R^m\times(0,+\infty)} \bar{\psi}_P(y, r). \tag{29}$$

The optimal value of (D_A) is denoted by M_A.

Denote by $\bar{S}$ the optimal set of (P). Let $\bar{V}(y, r)$ denote the optimal solution set of the problem

$$\bar{P}(y,r) \qquad \inf_{(x,u)\in R^n\times R^m} \left\{\bar{f}_P(x, u) + r\sigma(u) - \langle y, u\rangle\right\},$$

where $\bar{f}_P$ is defined by (3). Denote by $\bar{\psi}_P(y, r)$ the optimal value of $\bar{P}(y, r)$. Recall that M_P is the optimal value of problem (P).

From Theorem 2.1, we have the following result.

THEOREM 4.1. *Consider the constrained program* (P) *and its generalized augmented Lagrangian problem* $\bar{P}(y,r)$. *Suppose that*

$$\lim_{\|x\|\to+\infty, x\in X} \max\{f(x), g_1(x), \ldots, g_{m_1}(x), |g_{m_1+1}(x)|, \ldots, |g_m(x)|\} = +\infty, \tag{30}$$

and that $\bar{y}\in R^m$, $\bar{r}>0$ *satisfy*

$$\inf\{l_P(x,\bar{y},\bar{r}): x\in X\} > -\infty.$$

Then

(i) $\bar{S}$ *is nonempty and compact.*

(ii) *For any* $r\ge\bar{r}+1$, $\bar{V}(\bar{y},r)$ *is nonempty and compact.*

(iii) *For any selection* $(x(r),u(r))\in\bar{V}(y,r)$ *with* $r\ge\bar{r}+1$ *and* $r\to+\infty$, $\{(x(r),u(r))\}$ *is bounded, and its limit points take the form* $(x^*,0)$, *where* $x^*\in\bar{S}$.

(iv) $\lim_{r\to+\infty}\bar{\psi}_P(\bar{y},r)=M_P$.

(v) *Zero duality gap holds*:

$$M_A = M_P.$$

PROOF. Let us verify that all the conditions of Theorem 2.1 hold. It is obvious that the function φ defined by (2) is proper since the feasible set of the constrained program (P), $X_0\neq\varnothing$. The lsc of $\bar{f}_P$ is clear from the closedness of the set

$$\{(x,u)\in R^n\times R^m: \bar{f}_P(x,u)\le t\} = \{(x,u): x\in X, f(x)\le t, g_j(x)+u_j\le 0,$$
$$j=1,\ldots,m_1, g_j(x)+u_j=0, j=m_1+1,\ldots,m\},$$

for any $t\in R$ by the assumption that f, g_j $(j=1,\ldots,m_1)$ are lsc and g_j $(j=m_1+1,\ldots,m)$ are continuous and X is closed. Now we show that the condition (30) implies that $\bar{f}_P(x,u)$ defined by (3) is level-bounded in x locally uniform in u. Indeed, suppose to the contrary that there exist $\bar{\alpha}\in R$, $\bar{u}=(\bar{u}_1,\ldots,\bar{u}_m)\in R^m$, $u^k=(u_1^k,\ldots,u_m^k)\in R^m$ with

$$u^k\to\bar{u}, \tag{31}$$

and there exists $x^k\in R^n$ satisfying

$$\|x^k\|\to+\infty, \tag{32}$$

such that

$$\bar{f}_P(x^k,u^k)\le\bar{\alpha}, \qquad \forall k.$$

By the definition of $\bar{f}_P$, we see that, for any k,

$$\begin{gathered} x^k\in X,\\ f(x^k)\le\bar{\alpha},\\ g_j(x^k)+u_j^k\le 0, \qquad j=1,\ldots,m_1,\\ g_j(x^k)+u_j^k=0, \qquad j=m_1+1,\ldots,m. \end{gathered} \tag{33}$$

As a result,

$$\begin{aligned} &\max\{f(x^k), g_1(x^k), \ldots, g_{m_1}(x^k), |g_{m_1+1}(x^k)|, \ldots, |g_m(x^k)|\}\\ &\quad\le \max\{\bar{\alpha}, u_1^k, \ldots, u_{m_1+1}^k, |u_{m_1+1}^k|, \ldots, |u_m^k|\}\\ &\quad\le M \end{aligned}$$

for some $M>0$ by (31). It follows from (30) and (33) that $\{x^k\}$ is bounded. This contradicts (32). Thus, all the conditions of Theorem 2.1 are satisfied. The conclusions follow. $\square$

REMARK 4.1. From our assumption that $\inf_{x\in X} f(x) > 0$, we see that $l_P(x, 0, r) \geq 0$, $\forall\, x \in X$, $r > 0$ and as a result, $\bar{y}$ and $\bar{r}$ can take 0 and any $r \in (0, +\infty)$, respectively.

In the following, in the context of a constrained program (P), we establish an equivalence in terms of zero duality gap property between its generalized augmented Lagrangian dual scheme and its nonlinear Lagrangian dual scheme given in Rubinov et al. (1999b).

First we recall the definition of a nonlinear Lagrangian for the constrained program (P).

Let c: $R_+ \times R^{m_1} \times R_+^{m-m_1} \to R$ be a real-valued function. c is said to be *increasing* if, for any $y^1, y^2 \in R^{m_1} \times R_+^{m-m_1}$, $y^2 - y^1 \in R_+^{m+1}$ implies that $c(y^1) \leq c(y^2)$.

Consider increasing and lsc functions c defined on $R_+ \times R^{m_1} \times R_+^{m-m_1}$ having the following properties:

(A) $\exists\ a_j > 0, j = 1, \ldots, m$ such that, for any $y = (y_0, y_1, \ldots, y_{m_1}, y_{m_1+1}, \ldots, y_m) \in R_+ \times R^{m_1} \times R_+^{m-m_1}$, we have

$$c(y) \geq \max\left\{y_0, a_1 y_1, \ldots, a_{m_1} y_{m_1}, a_{m_1+1} y_{m_1+1}, \ldots, a_m y_m\right\}.$$

(B) For any $y_0 \in R_+$, we have $c(y_0, 0, \ldots, 0) = y_0$.

Let c be an increasing function with properties (A) and (B), and

$$F(x, d) = \left(f(x), d_1 g_1(x), \ldots, d_{m_1} g_{m_1}(x), d_{m_1+1}|g_{m_1+1}(x)|, \ldots, d_m |g_m(x)|\right),$$

where $d = (d_1, \ldots, d_m) \in R_+^m$ and $x \in X$.

The function defined by

$$L(x, d) = c(F(x, d)), \qquad x \in X, d \in R_+^m,$$

is called a *nonlinear Lagrangian* corresponding to c.

The *nonlinear Lagrangian dual function* corresponding to c is defined as

$$\phi(d) = \inf_{x\in X} L(x, d), \qquad d \in R_+^m.$$

The *nonlinear Lagrangian dual problem* is defined by

$$(D_N) \quad \sup_{d\in R_+^m} \phi(d). \tag{34}$$

Denote by M_N the optimal value of (D_N).

The following equivalence result extends Theorem 3.5 in Yang and Huang (2001) in the sense that the convexity and finiteness of the augmenting function (which implies that the augmenting function is locally Lipschitz) is replaced by the continuity of the generalized augmenting function at the origin. Since its proof is similar to that of Theorem 3.5 in Yang and Huang (2001), we omit the proof.

THEOREM 4.2. *Consider the constrained program* (P), *its generalized augmented Lagrangian dual problem* (D_A), *and its nonlinear Lagrangian dual problem* (D_N). *If the generalized augmenting function* σ *is continuous at* $0 \in R^m$ *and the increasing function* c *defining the nonlinear Lagrangian* L *is continuous, then the following two statements are equivalent*:

(i) $M_A = M_P$;

(ii) $M_N = M_P$.

4.2. Exact penalization. Now we apply Theorems 3.1 and 3.2 to the constrained program (P).

Let, for $u \in R^m$,

$$\bar{p}_1(u) = \inf\big\{f(x): x \in X, g_j(x) \le u_j, j = 1, \ldots, m_1, g_j(x) = u_j, j = m_1 + 1, \ldots, m\big\},$$

and

$$p_1(u) = \inf\{\bar{f}_P(x, u): x \in R^n\}, \qquad \forall u \in R^m,$$

where $\bar{f}_P(x, u)$ is defined by (3).

It is clear that $p_1(u) = \bar{p}_1(-u), \forall u \in R^m$.

DEFINITION 4.1. Consider the constrained program (P) and the associated generalized augmented Lagrangian $l_P(x, y, r)$. A vector $\bar{y} \in R^m$ is said to support an exact penalty representation if there exists $\bar{r} > 0$ such that

$$M_P = p_1(0) = \inf\{l_P(x, \bar{y}, r): x \in R^n\}, \qquad \forall r \ge \bar{r}$$

and

$$\arg\min\,(\mathrm{P}) = \arg\min_x l_P(x, \bar{y}, r), \qquad \forall r \ge \bar{r},$$

where $\arg\min$ (P) denotes the set of optimal solutions of (P).

THEOREM 4.3. *Consider the constrained program* (P). *The generalized augmented Lagrangian l is defined by* (4) *with the dualizing parameterization function $\bar{f}_P$ defined by* (3). *Then*

(i) *If $\bar{y}$ supports an exact penalty representation for* (P), *then there exist a neighborhood W of $0 \in R^m$ and $\bar{r} > 0$ such that*

$$p_1(u) \ge p_1(0) + \langle \bar{y}, u\rangle - \bar{r}\sigma(u), \qquad \forall u \in W,$$

which is equivalent to

$$\bar{p}_1(u) \ge \bar{p}_1(0) - \langle \bar{y}, u\rangle - \bar{r}\sigma(-u), \qquad \forall u \in W. \tag{35}$$

(ii) *The converse of* (i) *is true if $X_0 \ne \varnothing$, and there exist $\tau > 0$ and $N > 0$ such that $\sigma(u) \ge \tau\|u\|$ when $\|u\| \ge N$.*

PROOF. (i) is the direct consequence of statement (i) of Theorem 3.1.

By the assumption that $\inf_{x\in X} f(x) > 0$ and the condition in (ii), we see that

$$\begin{aligned}\bar{f}_P(x, u) - \langle \bar{y}, u\rangle + r\sigma(u) &\ge -\langle \bar{y}, u\rangle + r\sigma(u)\\ &\ge -\langle \bar{y}, u\rangle + \tau r\|u\| \ge (\tau r - \|\bar{y}\|)\|u\|,\end{aligned}$$

when $\|u\| \ge N$.

Taking $\bar{r}' = \|\bar{y}\|/\tau$, we know that $\bar{f}_P(x, u) - \langle \bar{y}, u\rangle + \bar{r}'\sigma(u)$ is bounded below.

Moreover, $X_0 \ne \varnothing$ guarantees that $p_1(0)$ is finite.

From (ii) of Theorem 3.1, statement (ii) of Theorem 4.3 follows. □

THEOREM 4.4. *With the notation and assumptions as in Theorem* 4.3, *we have*

(i) *If $\bar{y} = 0$ supports an exact penalty representation, then there exist a neighborhood W of $0 \in R^m$ and $\hat{r} > 0$ such that*

$$p_1(u) \ge p_1(0) - \hat{r}\sigma(u), \qquad \forall u \in W,$$

which is equivalent to

$$\bar{p}_1(u) \ge \bar{p}_1(0) - \hat{r}\sigma(-u), \qquad \forall u \in W. \tag{36}$$

(ii) *The converse of* (i) *is true if $X_0 \ne \varnothing$.*

PROOF. We only need to show that (ii) holds.

Since $\inf_{x\in X} f(x) > 0$, we see that

$$\bar{f}_P(x,u) + r\sigma(u) \geq 0, \qquad \forall\, x, u, r.$$

In addition, $X_0 \neq \varnothing$ is equivalent to $p_1(0) = \bar{p}_1(0)$ is finite. Thus, statement (ii) follows directly from Theorem 3.2(ii). $\square$

The following simple example illustrates that a generalized augmenting function may admit an exact penalty representation, while any augmenting function used in Rockafellar and Wets (1998) that is also finite near the origin $0 \in R^m$ cannot.

EXAMPLE 4.1. Consider the problem

$$\min -x \quad \text{s.t. } x \in R, \qquad x^3 \leq 0.$$

Clearly, $\bar{p}_1(u) = -u^{1/3}, \forall\, u \in R$. For any $\bar{y} \in R$, $\bar{r} > 0$, any neighborhood W of 0 and any augmenting function σ which is finite near $0 \in R$, consider the following inequality:

$$\bar{p}_1(u) \geq \bar{p}_1(0) - \bar{y}\cdot u - \bar{r}\sigma(-u), \qquad \forall\, u \in W,$$

i.e.,

$$-u^{1/3} \geq -\bar{y}\cdot u - \bar{r}\sigma(-u), \qquad \forall\, u \in W. \tag{37}$$

Since σ is convex and finite near $0 \in R$, there exists a constant $a > 0$ such that $\sigma(u) \leq a|u|$ when $|u|$ is sufficiently small. Thus,

$$-u^{1/3} \geq -\bar{y}\cdot u - \bar{r}a|u|,$$

when $|u|$ is sufficiently small. But this is impossible; i.e., (37) does not hold. Thus, if the augmenting function is finite near $0 \in R$, by Theorem 4.3(i), it is impossible to achieve an exact penalization representation. However, if we take $\sigma(u) = |u|^{1/3}$, $\bar{y} = 0$, $\bar{r} = 1$, then

$$\bar{p}_1(u) = -u^{1/3} \geq \bar{p}_1(0) - \bar{y}\cdot u - \bar{r}\sigma(-u) = -|u|^{1/3}, \qquad \forall\, u.$$

By Theorem 4.4, an exact penalty representation is obtained.

In the following, we apply some special cases of Theorem 4.4 to derive necessary and sufficient conditions for the classes of exact penalty functions $q_\gamma(x,r)$ and $L_\gamma(x,r)$ for the constrained program (P).

THEOREM 4.5. *Consider the constrained program* (P) *and the penalty function* $q_\gamma(x,r)$ *defined by* (5).

(i) *If there exists* $\bar{r} > 0$ *such that whenever* $r \geq \bar{r}$, $q_\gamma(x,r)$ *is an exact penalty function in the sense that*

$$M_P = \inf\{q_\gamma(x,r)\colon x \in X\} \tag{38}$$

and

$$\arg\min(P) = \arg\min_x q_\gamma(x,r), \tag{39}$$

then there exists a neighborhood W *of* $0 \in R^m$ *and* $\hat{r} > 0$ *such that*

$$\bar{p}_1(u) \geq \bar{p}_1(0) - \hat{r}\|u\|_1^\gamma, \qquad \forall\, u \in W.$$

(ii) *The converse of* (i) *is true if* $X_0 \neq \varnothing$.

REMARK 4.2. To the best of our knowledge, only some sufficient conditions have been provided in the literature for the exactness of the penalty function $q_\gamma(x, r)$. Theorem 4.5 gives a necessary and sufficient condition for the exactness of $q_\gamma(x, r)$.

Now we consider the class of nonlinear penalty functions $L_\gamma(x, d)$ defined by (6).

THEOREM 4.6. *Let $\gamma > 0$. Consider the constrained program* (P) *and the nonlinear penalty function $L_\gamma(x, r)$. Then the following two statements hold.*

(i) *if there exists $\bar{r} > 0$ such that when $r \geq \bar{r}$, $L_\gamma(x, r)$ is an exact penalty function in the sense that*

$$M_P = \inf\{L_\gamma(x, r)\colon x \in X\}, \tag{40}$$

and

$$\arg\min\ (P) = \arg\min_x L_\gamma(x, r), \tag{41}$$

then there exist a constant $M_3 > 0$ and a neighborhood W of $0 \in R^m$ such that

$$\bar{p}_1(u) \geq \bar{p}_1(0) - M_3 \sum_{j=1}^m |u_j|^\gamma, \qquad \forall\, u \in W. \tag{42}$$

(ii) *The converse of* (i) *is true if $X_0 \neq \varnothing$.*

PROOF. Let (P_γ) be as defined in Example 2.3. As noted in Example 2.3, problem (P_γ) is equivalent to problem (P) in the sense that the two problems have the same set of (local) solutions while the infimum of problem (P_γ) is exactly M_P^γ. Let l'_γ be defined as in Example 2.3.

Let

$$\theta_\gamma(x, r) = f^\gamma(x) + r\left[\sum_{j=1}^{m_1} g_j^{+\gamma}(x) + \sum_{j=m_1+1}^{m} |g_j(x)|^\gamma\right], \qquad x \in X, r > 0. \tag{43}$$

By Theorem 4.4, the following two statements are true:

(a) If there exists $\bar{r}' > 0$ such that whenever $r \geq \bar{r}'$, θ_γ is an exact penalty function for (P_γ) in the sense that

$$M_P^\gamma = \inf\{\theta_\gamma(x, r)\colon x \in X\} \tag{44}$$

and

$$\arg\min(P_\gamma) = \arg\min_x \theta_\gamma(x, r), \tag{45}$$

where $\arg\min(P_\gamma)$ is the set of optimal solutions of (P_γ), then there exist a constant $M_1 > 0$ and a neighborhood W of $0 \in R^m$ such that

$$\bar{p}_1^\gamma(u) \geq \bar{p}_1^\gamma(0) - M_1 \sum_{j=1}^m |u_j|^\gamma, \qquad \forall\, u \in W, \tag{46}$$

which is equivalent to the existence of a constant $M_2 > 0$ such that

$$\bar{p}_1(u) \geq \bar{p}_1(0) - M_2 \sum_{j=1}^m |u_j|^\gamma, \qquad \forall\, u \in W. \tag{47}$$

(b) The converse of (a) is true if $X_0 \neq \varnothing$.

Note that $\inf_{x\in R^n} \bar{f}_\gamma(x,u) = \bar{p}_1^\gamma(u), \forall u \in R^m$. The equivalence of (46) and (47) is shown as follows. First we show that (46) implies (47). By the mean-value theorem,

$$\bar{p}_1^\gamma(u) - \bar{p}_1^\gamma(0) = \gamma\xi^{\gamma-1}(\bar{p}_1(u) - \bar{p}_1(0)), \tag{48}$$

where ξ lies between $\bar{p}_1(u)$ and $\bar{p}_1(0)$. Note that if $\bar{p}_1(u) \geq \bar{p}_1(0)$, (47) holds automatically. Now assume that $\bar{p}_1(u) < \bar{p}_1(0)$. If $\gamma > 1$, then $\xi^{\gamma-1} \geq \bar{m}^{\gamma-1}$, where $\bar{m} = \inf_{x\in X} f(x) > 0$ by assumption. If $\gamma \in (0,1)$, then $\xi^{\gamma-1} \geq \bar{p}_1^{\gamma-1}(0) > 0$. Consequently, in either case, we always have $\xi^{\gamma-1} \geq m_0$ for some $m_0 > 0$. This, together with (46) and (48), proves (47). Similarly, we can prove that (47) implies (46) by noting $\bar{p}_1(u) = [\bar{p}_1^\gamma(u)]^{1/\gamma}, \forall u \in W$.

Finally, we observe that

$$\arg\min\,(\mathrm{P}) = \arg\min(P_\gamma), \qquad \forall\,\gamma > 0,$$

and the optimal value of (P_γ) is exactly M_P^γ. Note also that, for any $r > 0$,

$$\arg\min_x \theta_\gamma(x,r) = \arg\min_x L_\gamma(x,r),$$

and

$$\inf_{x\in X} \theta_\gamma(x,r) = \left[\inf_{x\in X} L_\gamma(x,r)\right]^\gamma.$$

These observations, together with statements (a) and (b), establish statements (i) and (ii). □

REMARK 4.3. Theorem 4.6 improves and extends the results in Theorem 7.1 in Rubinov et al. (1999a) where (41) was not obtained, and no equality constraint was considered.

4.3. Relationships among exact penalty functions. In what follows, we investigate relationships among the three types of exact penalty functions discussed in Theorems 4.3–4.6. The following lemma is useful.

LEMMA 4.1. *Suppose that $a_j \geq 0,\ j = 1,\ldots,m$. The following relations hold.*

(i) *If $\gamma \in (0,1]$, then*

$$\left(\sum_{j=1}^m a_j\right)^\gamma \leq \sum_{j=1}^m a_j^\gamma \leq m\left(\sum_{j=1}^m a_j\right)^\gamma. \tag{49}$$

(ii) *If $\gamma \in [1,+\infty)$, then*

$$\frac{1}{m^{\gamma-1}}\left(\sum_{j=1}^m a_j\right)^\gamma \leq \sum_{j=1}^m a_j^\gamma \leq \left(\sum_{j=1}^m a_j\right)^\gamma. \tag{50}$$

PROOF. Consider the function $f_1(t,\gamma) = t^\gamma$ defined on $[0,1]\times(0,+\infty)$. It is clear that f_1 is nonincreasing with respect to γ on $(0,+\infty)$ for any fixed $t \in [0,1]$. It can also be checked that f_1 is convex in t for any fixed $\gamma \in [1,+\infty)$. If $a_i = 0, i = 1,\ldots,m$, then (49) and (50) hold automatically. Thus, we assume, without loss of generality, that there exists at least one $a_i > 0$.

(i) If $\gamma \in (0,1]$, it follows from the nonincreasing property in γ of f_1 for any fixed $t \in [0,1]$ that

$$\frac{\sum_{j=1}^m a_j^\gamma}{(\sum_{j=1}^m a_j)^\gamma} = \sum_{j=1}^m \left(\frac{a_j}{\sum_{j=1}^m a_j}\right)^\gamma \geq \sum_{j=1}^m \frac{a_j}{\sum_{j=1}^m a_j} = 1. \tag{51}$$

Moreover, it is obvious that

$$\frac{\sum_{j=1}^m a_j^\gamma}{(\sum_{j=1}^m a_j)^\gamma} \le m, \tag{52}$$

since $a_j^\gamma \le (\sum_{j=1}^m a_j)^\gamma$, $j = 1, \ldots, m$. Relation (49) follows immediately from (51) and (52).

(ii) If $\gamma \in [1, +\infty)$, then

$$\frac{\sum_{j=1}^m a_j^\gamma}{(\sum_{j=1}^m a_j)^\gamma} = \sum_{j=1}^m \left(\frac{a_j}{\sum_{j=1}^m a_j}\right)^\gamma \le \sum_{j=1}^m \frac{a_j}{\sum_{j=1}^m a_j} = 1. \tag{53}$$

By the convexity of f_1 in t for any fixed $\gamma \in [1, +\infty)$, we have

$$\begin{aligned} \frac{\sum_{j=1}^m a_j^\gamma}{(\sum_{j=1}^m a_j)^\gamma} &= \sum_{j=1}^m \left(\frac{a_j}{\sum_{j=1}^m a_j}\right)^\gamma \\ &= m\sum_{j=1}^m 1/m\left(\frac{a_j}{\sum_{j=1}^m a_j}\right)^\gamma \\ &\ge m\left[\sum_{j=1}^m 1/m\frac{a_j}{\sum_{j=1}^m a_j}\right]^\gamma \\ &= \frac{1}{m^{\gamma-1}}. \end{aligned} \tag{54}$$

(50) follows directly from (53) and (54). □

The following result shows that the penalty functions $q_\gamma(x, r)$ and $L_\gamma(x, r)$ are equivalent in the sense that if one of them is exact, then the other is also exact when the penalty parameter is large enough.

THEOREM 4.7. *Let $\gamma > 0$. Consider the constrained program* (P), *the penalty functions $q_\gamma(x, r)$ and $L_\gamma(x, r)$. Suppose that $X_0 \neq \varnothing$. Then the following statements are equivalent.*

(i) *There exists $\bar{r} > 0$ such that $L_\gamma(x, r)$ is an exact penalty function whenever $r \ge \bar{r}$.*

(ii) *There exists $\bar{r} > 0$ such that $q_\gamma(x, r)$ is an exact penalty function whenever $r \ge \bar{r}$.*

PROOF. The conclusion follows directly from Theorems 4.5, 4.6, and Lemma 4.1. □

The next theorem establishes that if the generalized augmented Lagrangian $l_P(x, y, r)$ admits an exact penalty representation, then the penalty functions $q_\gamma(x, r)$ and $L_\gamma(x, r)$ $(0 < \gamma \le 1)$ are both exact when the penalty parameter is sufficiently large.

THEOREM 4.8. *Consider the constrained optimization problem* (P), *its associated generalized augmented Lagrangian $l(x, y, r)$ defined in* (4) *with the dualizing parameterization function $\bar{f}_P(x, u)$ defined by* (3), *the penalty functions $q_\gamma(x, r)$ defined by* (5), *and $L_\gamma(x, r)$ defined by* (6), *respectively. Assume that $X_0 \neq \varnothing$. Suppose that the generalized augmenting function σ is locally Lipschitz at the origin $0 \in R^m$ and $\bar{y}$ supports an exact penalty representation. Then*

(a) *For any $\gamma \in (0, 1]$, there exists $\bar{r} > 0$ such that $L_\gamma(x, r)$ is an exact penalty function whenever $r \ge \bar{r}$.*

(b) *For any $\gamma \in (0, 1]$, there exists $\bar{r} > 0$ such that $q_\gamma(x, r)$ is an exact penalty function whenever $r \ge \bar{r}$.*

PROOF. Since the generalized augmenting function σ is locally Lipschitz at the origin $0 \in R_+^m$, there exist $L > 0$ and $\theta > 0$ such that

$$0 \le \sigma(u) \le L\|u\|_1, \qquad \forall\, u \in R^m \text{ with } \|u\|_1 \le \theta < 1. \tag{55}$$

As $\bar{y}$ supports an exact penalty representation, by Theorem 4.3, there exist $\hat{r} > 0$ and a neighborhood W of the origin $0 \in R^m$ such that (35) holds. Without loss of generality, we assume that $W = \{u \in R^m\colon \|u\|_1 \le \theta\}$. The combination of (35) and (55) yields

$$\begin{aligned} \bar{p}_1(u) &\ge \bar{p}_1(0) - \langle \bar{y}, u\rangle - \hat{r}\|u\|_1 \\ &\ge \bar{p}_1(0) - \|\bar{y}\|_1\|u\|_1 - \hat{r}\|u\|_1 \\ &= \bar{p}_1(0) - (\|\bar{y}\|_1 + \hat{r})\|u\|_1, \qquad \forall\, u \in W. \end{aligned} \tag{56}$$

Observing that when $0 < \gamma \le 1$, we have

$$\|u\|_1 \le \sum_{j=1}^{m} |u_j|^{\gamma}. \tag{57}$$

(56) and (57) imply

$$\bar{p}_1(u) \ge \bar{p}_1(0) - (\|\bar{y}\|_1 + \hat{r})\sum_{j=1}^{m}|u_j|^{\gamma}, \qquad \forall\, u \in W.$$

Thus (a) holds by Theorem 4.6.

Statement (b) follows from Theorem 4.7 and statement (a). □

REMARK 4.4. Any augmenting function σ, which is finite near $0 \in R^m$, is locally Lipschitz at $0 \in R^m$.

Finally, for a constrained program (P), we establish an equivalence in terms of exact penalization among the generalized augmented Lagrangian $l_P(x, y, r)$, the penalty functions $q_\gamma(x, r)$ and $L_\gamma(x, r)$.

THEOREM 4.9. *Assume the same conditions as Theorem* 4.8. *Let* $\gamma > 0$. *Suppose that the generalized augmenting function* σ *satisfies*

$$0 < \liminf_{u\to 0} \sigma(u)/\|u\|_1^{\gamma} \le \limsup_{u\to 0} \sigma(u)/\|u\|_1^{\gamma} < +\infty. \tag{58}$$

Then the following three statements are equivalent:

(a) $\bar{y} = 0$ *supports an exact penalty representation in the framework of the generalized augmented Lagrangian* l.

(b) *There exists* $\bar{r} > 0$ *such that* $q_\gamma(x, r)$ *is an exact penalty function whenever* $r \ge \bar{r}$.

(c) *There exists* $\bar{r} > 0$ *such that* $L_\gamma(x, r)$ *is an exact penalty function whenever* $r \ge \bar{r}$.

PROOF. By Theorem 4.4, $\bar{y} = 0$ supports an exact penalty representation if and only if there exist $\hat{r} > 0$ and a neighborhood W of the origin $0 \in R^m$ such that

$$\bar{p}_1(u) \ge \bar{p}_1(0) - \hat{r}\sigma(u), \qquad \forall\, u \in W. \tag{59}$$

Without loss of generality, assume that $W = \{u \in R^m\colon \|u\|_1 \le \theta < 1\}$, where $\theta > 0$.

Since (58) holds, it follows that (59) holds if and only if there exist a scalar $\eta > 0$ and a neighborhood W' of $0 \in R^m$ such that

$$\bar{p}_1(u) \ge \bar{p}_1(0) - \eta\|u\|_1^{\gamma}, \qquad \forall\, u \in W'. \tag{60}$$

Applying Theorem 4.5, we see that (60) holds if and only if there exists $\bar{r} > 0$ such that (38) and (39) hold. This proves that (a) and (b) are equivalent. The equivalence between (b) and (c) is given in Theorem 4.7. □

Acknowledgments. The authors are very grateful to the Area Editor, Professor Jong-Shi Pang, an Associate Editor, and two referees, whose inspiring criticisms and constructive comments have significantly improved the presentation of the paper. This research is supported by the Research Grants Council of Hong Kong (B-Q359).

References

Auslender, A. 1999. Penalty and barrier methods: A unified framework. *SIAM J. Optim.* **10** 211–230.

———, R. Cominetti, M. Haddou. 1997. Asymptotical analysis for penalty and barrier methods in convex and linear programming. *Math. Oper. Res.* **22** 43–62.

Bestsekas, D. P. 1982. *Constrained Optimization and Lagrangian Multiplier Methods.* Academic Press, New York.

Burke, J. V. 1991a. Calmness and exact penalization. *SIAM J. Control and Optim.* **29** 493–497.

———. 1991b. An exact penalization viewpoint of constrained optimization. *SIAM J. Control and Optim.* **29** 968–998.

Huang, X. X., X. Q. Yang. 2001. Duality and exact penalization for vector optimization via augmented Lagrangian. *J. Optim. Theory Appl.* **111** 615–640.

———, ———. Generalized augmented Lagrangian methods for equality constrained optimization problems. Working paper, Hong Kong Polytechnic University, Hong Kong, China.

Ioffe, A. 1979. Necessary and sufficient conditions for a local minimum. 3: Second-order conditions and augmented duality. *SIAM J. Control and Optim.* **17** 266–288.

Luo, Z. Q., J. S. Pang. Eds. 2000. Error bounds in mathematical programming. *Math. Programming Ser. B.* **88**(2).

———, ———, D. Ralph. 1996. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, New York.

Pang, J. S. 1997. Error bounds in mathematical programming. *Math. Programming* **79** 299–332.

Rockafellar, R. T. 1974. Augmented Lagrange multiplier functions and duality in nonconvex programming. *SIAM J. Control and Optim.* **12** 268–285.

———. 1993. Lagrange multipliers and optimality. *SIAM Rev.* **35** 183–238.

———, R. J.-B. Wets. 1998. *Variational Analysis*. Springer-Verlag, Berlin.

Rubinov, A. M., B. M. Glover, X. Q. Yang. 1999a. Decreasing functions with applications to penalization. *SIAM J. Optim.* **10**(1) 289–313.

———, ———, ———. 1999b. Modified Lagrangian and penalty functions in continuous optimization. *Optim.* **46** 327–351.

Yang, X. Q., X. X. Huang. 2001. A nonlinear Lagrangian approach to constrained optimization problems. *SIAM J. Optim.* **11** 1119–1144.

X. X. Huang: Department of Mathematics and Computer Science, Chongqing Normal University, Chongqing 400047, China; e-mail: mahuangx@polyu.edu.hk

X. Q. Yang: Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China; e-mail: mayangxq@polyu.edu.hk