Journal of Global Optimization **21:** 51–65, 2001.

Approximate Optimal Solutions and Nonlinear Lagrangian Functions*

X.X. HUANG[1,2] and X.Q. YANG[3]
[1]*Department of Mathematics and Computer Science, Chongqing Normal University, Chongqing 400047, China.* [2]*Current address: Department of Applied Mathematics, Hong Kong Polytechnic University, Kowloon, Hong Kong;* [3]*Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong (e-mail: mayangxq@polyu.edu.hk)*

(Received 6 June 2000; accepted in revised form 22 March 2001)

Abstract. There is an increasing interest in the study of optimality conditions of approximate solutions for nonlinear optimization problems. In this paper, relationships between approximate optimal values and approximate roots of a nonlinear function are explored via a nonlinear Lagrangian function. Almost approximate optimal solutions are investigated by means of nonlinear Lagrangian functions.

Key words: Nonlinear programming, Approximate solution, Nonlinear Lagrangian, Ekeland's variational principle

1. Introduction

Unconstrained optimization technique is a popular method in solving nonlinear optimization problems, such as dual methods and penalty function methods. Many of these methods are established based on the assumption that an optimal solution of the nonlinear optimization problem exists. However, it is well-known that many mathematical programs do not have an optimal solution and thus we have to resort to approximate solutions ([2,8,9]). Moreover, sometimes we do not need to find an exact optimal solution even if it does exist due to the fact that it is often very expensive to find an exact solution. As a matter of fact, most numerical methods attempting at global optimization only yield approximate optimal solutions. There have been some works in the literature which are devoted to investigating approximate solutions of the constrained optimization problems by means of penalty functions (see, e.g., [8,9,11,19,20]).

Recently a general class of nonconvex constrained optimization problems have been reformulated as unconstrained optimization problems via nonlinear Lagrangians. Under some conditions, necessary and sufficient optimality conditions, duality theory, saddle point theory as well as exact penalization results among the original

* This work is supported by a grant from the Australia Research Council and it was done while the first author was visiting the University of Western Australia.

constrained optimization and its unconstrained nonlinear Lagrangian problem have been established (see, e.g., [3,13,14,15,17,18]). Based on the results obtained in Goh and Yang [3], Li and Sun [6] developed several algorithms for solving a constrained optimization problem by seeking the smallest root of a nonlinear function of a real variable. It is worth noting that most of these results are established on the basis of the assumption that the set of optimal solutions of the original constrained optimization problem is not empty. So it is interesting to study approximate solutions of a constrained mathematical program which does not necessarily have an optimal solution by means of nonlinear Lagrangian functions. It is well-known that Ekeland's variational principle and penalty functions are effective tools for studying approximate optimization (see, e.g., [2,7–11,16,19,20]). In particular, we would like to mention their effectiveness in the study of necessary optimality conditions for constrained and unconstrained programs and in the development of numerical methods for approximate optimization. Nonlinear Lagrangian functions have some similar properties of penalty functions. So it is possible to apply them in the study of approximate solutions of constrained optimization problems.

In this paper we develop approximate methods for solving a class of nonconvex constrained optimization problems by finding approximate roots of a well-behaved nonlinear function. This can be seen as a further development of the theory and method proposed in [3], [15] and [17]. We also investigate the possibility of obtaining the various versions of approximate solutions to a constrained program by solving an unconstrained program formulated by using a general nonlinear Lagrangian function. These results would be useful for the development of numerical methods for a constrained program by solving an unconstrained program and for the derivation of optimality conditions for various approximate solutions to constrained optimization problems. As an application, a KKT type optimality condition is obtained for a kind of approximate solution to a constrained program. Our results are applicable to nonconvex Lipschitz programming problems that satisfy a Mangasarian-Fromovitz constraint qualification.

This paper is organized as follows. In Section 2, we present some concepts, basic assumptions and preliminary results. Section 3 deals with the relationship between approximate optimal values and approximate roots of a nonlinear function which is constructed via a type of nonlinear Lagrangian function. Section 4 studies the approximate optimal solutions by means of the approximate optimal solutions of the nonlinear Lagrangian functions. Section 5 concludes the paper.

2. Preliminaries

In this section, we introduce some definitions and Ekeland's variational principle. Consider the following constrained optimization problem:

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

where $X \subseteq R^n$ is a nonempty and closed set, $f : X \to R$, $g_j : X \to R$, $j = 1, ..., m$ are lower semi continuous (l.s.c. in short) functions. Let M_P denote the optimal value of (P).

In this paper, we assume, without loss of generality, that

$$\inf_{x \in X} f(x) > 0.$$

Otherwise, we may replace the objective function $f(x)$ by $1 + e^{f(x)}$ and the resulting constrained optimization problem has the same optimal solutions as that of (P).

We denote by X_0 the set of feasible solutions, i.e.,

$$X_0 = \{x \in X : g_j(x) \le 0, \quad j = 1, ..., m\}.$$

Throughout the paper, we always assume that $X_0 \neq \phi$. For any $\epsilon > 0$, we denote by $X(\epsilon)$ the set of ϵ feasible solutions, i.e.,

$$X(\epsilon) = \{x \in X : g_j(x) \le \epsilon, \quad j = 1, ..., m\}.$$

For the sake of convenience, we set

$$g(x) = \max\{g_1(x), \ldots, g_m(x)\}, \quad \forall x \in X.$$

The following various definitions of approximate solutions are taken from Loridan [9].

DEFINITION 2.1 *Let $\epsilon > 0$. The point $x^* \in X_0$ is called an ϵ-solution of (P) if*

$$f(x^*) \le f(x) + \epsilon, \quad \forall x \in X_0.$$

DEFINITION 2.2 *Let $\epsilon > 0$. The point $x^* \in X_0$ is called an ϵ-quasi solution of (P) if*

$$f(x^*) \le f(x) + \epsilon \|x - x^*\|, \quad \forall x \in X_0.$$

REMARK 2.1 *An ϵ-quasi solution is also a locally ϵ-solution. In fact, x^* is an ϵ-solution of f on $\{x \in X_0 : \|x - x^*\| \le 1\}$.*

DEFINITION 2.3 *Let* $\epsilon > 0$. *If* $x^* \in X_0$ *is both an* ϵ*-solution and an* ϵ*-quasi solution of (P), we say that* x^* *is a regular* ϵ*-solution of (P).*

DEFINITION 2.4 *Let* $\epsilon > 0$. *If* $x^* \in X(\epsilon)$ *and*

$$f(x^*) \leq f(x) + \epsilon, \forall x \in X_0,$$

we say that x^* *is an almost* ϵ*-solution of (P).*

DEFINITION 2.5 *The point* $x^* \in X$ *is said to be an almost regular* ϵ*-solution of (P) if*

(i) $x^* \in X(\epsilon)$;
(ii) $f(x^*) \leq f(x) + \epsilon, \quad \forall x \in X_0$;
(iii) $f(x^*) \leq f(x) + \epsilon \|x - x^*\|, \quad \forall x \in X_0$.

PROPOSITION 2.1 *[2]. (Ekeland's variational principle). For any* $\epsilon > 0$, *there exists an* $x^* \in X_0$ *such that*

(i) $f(x^*) \leq f(x) + \epsilon, \quad \forall x \in X_0$;
(ii) $f(x^*) < f(x) + \epsilon \|x - x^*\|, \quad \forall x \in X_0\{x^*\}$.

Consequently, x^* *is a regular* ϵ*-solution of (P).*

3. Approximate Optimal Solutions, Approximate Optimal Values and Approximate Roots of A Nonlinear Function

In this section, we shall firstly establish approximate necessary and sufficient conditions for (P). These results are parallel to that of Theorem 1 in [3], but different forms of constrained optimization problems and different forms of unconstrained programs are considered. Secondly, we will define a nonlinear function ϕ_ϵ with a precision parameter $\epsilon \geq 0$ and explore the relationship between the approximate optimal values of (P) and the approximate roots of ϕ_ϵ.

Consider the following unconstrained optimization problem:

$$\text{(Q)} \qquad \inf_{x \in X} \max\{f(x) - M_P, g(x)\}.$$

PROPOSITION 3.1 *We have*

(i) If x^* *is an* ϵ*-solution of (P), then* x^* *is an* ϵ*-solution of (Q).*
(ii) If x^* *is an* ϵ*-solution of (Q), then* x^* *is an almost* ϵ*-solution of (P).*

Proof. (i) Suppose that x^* is an ϵ-solution of (P), then

$$f(x^*) \le M_P + \epsilon, \quad \forall x \in X_0.$$

To prove that x^* is an ϵ-solution of (Q), we only need to show that

$$\begin{aligned} f(x^*) - M_P &= \max\{f(x^*) - M_P, g(x^*)\} \\ &\le \max\{f(x) - M_P, g(x)\} + \epsilon, \quad \forall x \in X. \end{aligned} \tag{3.1}$$

When $x \in X_0$, it is sufficient to show that

$$f(x^*) - M_P \le f(x) - M_P + \epsilon.$$

This is obvious since x^* is an ϵ-solution of (P). If $x \notin X_0$, then $g(x) > 0$. To prove (3.1), it is sufficient to show that

$$f(x^*) \le M_P + \epsilon.$$

This follows from the fact that x^* is an ϵ-solution of (P).

(ii) Suppose that x^* is an ϵ-solution of (Q). Then

$$\max\{f(x^*) - M_P, g(x^*)\} \le \max\{f(x) - M_P, g(x)\} + \epsilon, \forall x \in X.$$

In particular, we have

$$\max\{f(x^*) - M_P, g(x^*)\} \le \inf_{x \in X_0} \max\{f(x) - M_P, g(x)\} + \epsilon = 0 + \epsilon = \epsilon.$$

Hence,

$$f(x^*) \le M_P + \epsilon,$$

$$g_j(x^*) \le \epsilon, \ j = 1, ..., m.$$

□

REMARK 3.1 *In Proposition 3.1, if ϵ reduces to 0, then the ϵ-solutions of (P) and (Q) are exact solutions of (P) and (Q), respectively; an almost ϵ-solution of (P) is also an exact solution of (P). In this case, solving (P) is equivalent to solving (Q) in the sense that the two problems have the same sets of optimal solutions.*

Let

$$\phi_\epsilon(\theta) = \inf_{x \in X} \max\{f(x) - \theta + \epsilon, g(x)\}, \quad \epsilon \ge 0, \theta \in R.$$

It is obvious that when $\epsilon = 0$, seeking the value of $\phi_\epsilon(M_P)$ involves solving problem (Q). From Remark 3.1, we know that solving (Q) will further give us the solution of problem (P). This observation means that there may exist some close relationship between the function ϕ_ϵ and the solution of problem (P). In the following we shall investigate this relationship.

First we state the following properties of ϕ_ϵ, which can be easily checked.

PROPOSITION 3.2 *We have*

(i) $\phi_\epsilon(M_P) \geq 0, \forall \epsilon \geq 0$; $\varphi_\epsilon(M_P - \epsilon) = \varphi_0(M_P), \forall \epsilon \geq 0$;
(ii) ϕ_ϵ *is a continuous, and non increasing function of* θ;
(iii) *for each fixed* θ, $\phi_\epsilon(\theta)$ *is nondecreasing and continuous with respect to* ϵ;
(iv) $\phi_\epsilon(\theta)$ *is continuous in* (θ, ϵ).

PROPOSITION 3.3 *For any* $\epsilon > 0$, *there exists* $x_\epsilon \in X_0$ *such that* x_ϵ *is an* ϵ-*solution of (P) and*

$$0 \leq \phi_\epsilon(f(x_\epsilon)) \leq \epsilon. \tag{3.2}$$

Proof. It is clear that for any $\epsilon > 0$, there exists $x_\epsilon \in X_0$ such that

$$f(x_\epsilon) \leq f(x) + \epsilon, \quad \forall x \in X_0.$$

That is,

$$f(x) - f(x_\epsilon) + \epsilon \geq 0, \quad \forall x \in X_0.$$

Therefore,

$$\max\{f(x) - f(x_\epsilon) + \epsilon, g(x)\} \geq 0, \quad \forall x \in X_0. \tag{3.3}$$

In addition,

$$\max\{f(x) - f(x_\epsilon) + \epsilon, g(x)\} \geq 0, \quad \forall x \in X \backslash X_0. \tag{3.4}$$

It follows from (3.3) and (3.4) that

$$\phi_\epsilon(f(x_\epsilon)) \geq 0.$$

Furthermore,

$$\phi_\epsilon(f(x_\epsilon)) \leq \max\{f(x_\epsilon) - f(x_\epsilon) + \epsilon, g(x_\epsilon)\} = \epsilon.$$

Thus, we have completed the proof. □

REMARK 3.2 *1. It is evident from the proof that for any* ϵ-*solution* x_ϵ *of (P), (3.2) holds.*

2. $a \in R^1$ *is called an* ϵ-*optimal value of (P) if there exists* $x \in X_0$ *such that* $a = f(x)$ *and* $a \leq M_p + \epsilon$. a *is called an* ϵ-*root of* ϕ_ϵ *if* $0 \leq \phi_\epsilon(a) \leq \epsilon$. *Thus, Proposition 3.3 implies that any* ϵ-*optimal value of (P) is an* ϵ-*root of* ϕ_ϵ.

Let $\epsilon \geq 0$. Consider the perturbed problem of (P):

$$(P_\epsilon) \qquad \begin{array}{ll} \inf & f(x) \\ \text{s.t.} & x \in X \\ & g_j(x) \leq \epsilon,\ j = 1, \cdots, m. \end{array}$$

Let $\beta(\epsilon)$ denote the infimum of (P_ϵ). Clearly, $\beta(0) = M_P$.

PROPOSITION 3.4 *We have $\phi_0(M_P) = 0$. Furthermore, if*

$$\lim_{\epsilon \to 0^+} \beta(\epsilon) = M_P, \tag{3.5}$$

then M_P is the smallest root of ϕ_0, where $\epsilon \to 0^+$ means that $\epsilon > 0$ and $\epsilon \to 0$.

Proof. By Proposition 3.3, for any $\epsilon > 0$, there exists $x_\epsilon \in X_0$ such that (3.2) and the following relation hold.

$$M_P \leq f(x_\epsilon) \leq M_P + \epsilon. \tag{3.6}$$

Taking limit in (3.2) when $\epsilon \to 0^+$ and applying (iv) of Proposition 3.2 as well as (3.6), we get $\phi_0(M_P) = 0$. Now assume that (3.5) holds. Let us show by contradiction that M_P is the smallest root of ϕ_0.

Suppose to the contrary that there exists $\delta > 0$ and

$$\theta^* \leq M_P - \delta \tag{3.7}$$

such that $\phi_0(\theta^*) = 0$. Then there exists $x_k \in X$ and $\epsilon_k \downarrow 0$ such that

$$\max\{f(x_k) - \theta^*, g(x_k)\} \leq \epsilon_k.$$

That is,

$$f(x_k) - \theta^* \leq \epsilon_k \tag{3.8}$$

and

$$g_j(x_k) \leq \epsilon_k, \; j = 1, \ldots, m. \tag{3.9}$$

It follows from (3.9) that $\beta(\epsilon_k) \leq f(x_k)$. This combined with (3.8) yields

$$\beta(\epsilon_k) \leq f(x_k) \leq \theta^* + \epsilon_k. \tag{3.10}$$

It follows from (3.7) and (3.10) that

$$\beta(\epsilon_k) \leq M_P - \delta + \epsilon_k. \tag{3.11}$$

Letting $k \to +\infty$ in (3.11), we obtain

$$\liminf_{k \to +\infty} \beta(\epsilon_k) \leq M_P - \delta,$$

which contradicts (3.5). □

REMARK 3.3 *1. Proposition 3.4 shows that the optimal value of (P) is a root of ϕ_0. When (3.5) holds, the optimal value of (P) is the least root of ϕ_0. The algorithm suggested in [17] can be applied to seek the least root of ϕ_0.*

2. Any one of the following conditions can guarantee (3.5) (see [4,14]):

(i) X is nonempty and bounded;
(ii) if X is unbounded, $\lim_{\|x\| \to +\infty, x \in X} f(x) = +\infty$;
(iii) if X is unbounded, there exist $\alpha > 0$ *and* $N > 0$ *such that* $g(x) \geq \alpha, \forall x \in X$ *such that* $\|x\| > N$;
(iv) if X is unbounded, $\lim_{\|x\| \to +\infty, x \in X} \max\{f(x), g(x)\} = +\infty$;
(v) the set-valued map $X(\epsilon)$ *is upper semi continuous (u.s.c. in short) at 0 (for example,* $\exists \epsilon_0 > 0$ *such that* $X(\epsilon_0)$ *is nonempty and compact) and* f *is uniformly continuous on a neighborhood* U *of* X_0.

3. Compared with Theorem 7.2 in [15], Proposition 3.4 has the following advantages:

(a) we did not assume the existence of an optimal solution for (P);
(b) we only assumed that f, g_j *are l.s.c. rather than continuous;*
(c) our assumption (3.5) is much weaker than the conditions (i) and (ii) listed above which were used in Theorem 7.2 in [15].

PROPOSITION 3.5 *Let* $X_1 = \{x \in X : g_j(x) < 0, j = 1, ..., m\}$. *If* $0 \leq \phi_\epsilon(\theta_\epsilon) \leq \epsilon$, *then for any* $\epsilon' > \epsilon$, $\exists x_{\epsilon'} \in X$ *such that*

$$f(x_{\epsilon'}) \leq \theta_\epsilon + \epsilon' - \epsilon, \tag{3.12}$$

$$g_j(x_{\epsilon'}) \leq \epsilon', \; j = 1, ..., m. \tag{3.13}$$

Furthermore, if either of the following two conditions holds:
(C_1) *for any* $x \in X_0$ *such that* $f(x)+\epsilon < \theta_\epsilon$, $\exists x' \in X_1$ *such that* $f(x')+\epsilon < \theta_\epsilon$;
(C_2) *there exists* $x_k \in X_1, k = 1, 2, ...$ *such that* $f(x_k) \to M_P$, *then*

$$\theta_\epsilon \leq f(x) + \epsilon, \quad \forall x \in X_0. \tag{3.14}$$

Proof. It follows from $\phi_\epsilon(\theta_\epsilon) \leq \epsilon$ and the definition of ϕ_ϵ that for any $\epsilon' > \epsilon$, $\exists x_{\epsilon'} \in X$ such that

$$\max\{f(x_{\epsilon'}) - \theta_\epsilon + \epsilon, g(x_{\epsilon'})\} \leq \epsilon', \quad \forall x \in X.$$

That is,

$$f(x_{\epsilon'}) \leq \theta_\epsilon + \epsilon' - \epsilon,$$

$$g_j(x_{\epsilon'}) \leq \epsilon', \; j = 1, ..., m.$$

Thus, (3.12) and (3.13) hold. In what follows, we prove that if (C_1) or (C_2) holds, then (3.14) holds. Suppose that (C_1) holds. Let $x \in X_0$. From $\phi_\epsilon(\theta_\epsilon) \geq 0$, we deduce that

$$\max\{f(x) - \theta_\epsilon + \epsilon, g(x)\} \geq 0, \forall x \in X_0. \tag{3.15}$$

If $\max\{f(x)-\theta_\epsilon+\epsilon, g(x)\} > 0$, then $f(x)-\theta_\epsilon+\epsilon > 0$. Namely, $\theta_\epsilon \le f(x)+\epsilon$. If $\max\{f(x)-\theta_\epsilon+\epsilon, g(x)\} = 0$, we show by contradiction that $\theta_\epsilon \le f(x)+\epsilon$. Otherwise, $f(x)+\epsilon < \theta_\epsilon$, by condition (C_1), there exists $x' \in X_1 \subset X_0$ such that $f(x')+\epsilon < \theta_\epsilon$. Consequently, $\max\{f(x')-\theta_\epsilon+\epsilon, g(x')\} < 0$, contradicting (3.15). So (3.14) has been proved.

Now suppose that (C_2) holds. Suppose to the contrary that (3.14) does not hold. Then $\theta_\epsilon > M_P+\epsilon$. By condition (C_2), we have $x_k \in X_0$ with

$$g(x_k) < 0 \tag{3.16}$$

such that $f(x_k) \to M_P$. Thus,

$$\theta_\epsilon > f(x_k)+\epsilon \tag{3.17}$$

when k is sufficiently large. Moreover, it follows from (3.15) and $x_k \in X_0$ that

$$\max\{f(x_k)-\theta_\epsilon+\epsilon, g(x_k)\} \ge 0, \quad \forall x \in X_0. \tag{3.18}$$

The combination of (3.16), (3.17) and (3.18) yields a contradiction. So (3.14) holds.
□

REMARK 3.4 *1. (3.12), (3.13) and (3.14) together show that $x_{\epsilon'}$ is an ϵ'-almost solution of (P).*

2. Proposition 3.3 says that an ϵ-optimal value is an ϵ-root of ϕ_ϵ. However, the converse may not be true even under condition (C_1) or (C_2). But under condition (C_1) or (C_2), an ϵ-root of ϕ_ϵ is indeed an almost approximate optimal value of (P), that is, for any $\epsilon' > \epsilon$, $\exists x_{\epsilon'} \in X$ such that (3.12)–(3.14) hold.. To obtain an ϵ-root of ϕ_ϵ. One can use the bisection method by the following procedure:

Step 1: Choose $\theta^1, \theta^2 \in R^1$ such that $\phi_\epsilon(\theta^1) \ge 0$ and $\phi_\epsilon(\theta^2) \le 0$.

Step 2: If $\phi_\epsilon(\theta^1) \le \epsilon$, then set $\theta_\epsilon = \theta^1$ and stop; if $\phi_\epsilon(\theta^2) = 0$, then set $\theta_\epsilon = \theta^2$ and stop. Otherwise, go to Step 3.

Step 3: Compute $\phi_\epsilon(\frac{\theta^1+\theta^2}{2})$ via any global minimizing algorithm. If $0 \le \phi_\epsilon(\frac{\theta^1+\theta^2}{2}) \le \epsilon$, then set $\theta_\epsilon = \frac{\theta^1+\theta^2}{2}$ and stop. Otherwise, go to Step 4.

Step 4: If $\phi_\epsilon(\frac{\theta^1+\theta^2}{2}) < 0$, then set $\theta^2 = \frac{\theta^1+\theta^2}{2}$ and go to Step 3. If $\phi_\epsilon(\frac{\theta^1+\theta^2}{2}) > \epsilon$, then set $\theta^1 = \frac{\theta^1+\theta^2}{2}$ and go to Step 3.

3. It is not difficult to verify that if $X_0 = clX_1$ (‘cl’ stands for the closure of a set) and f is continuous, then condition (C1) holds.

4. If either of the following conditions holds, then both (C_1) and (C_2) hold:

(i) $f, g_j (j = 1, ..., m)$ are locally Lipschitz, and the generalized Mangasarian-Fromovitz constraint qualification holds at each $x \in X_0$: there exists $u \in T_X(x)$ such that

$$\limsup_{t\to 0} \frac{g_j(x+tu)-g_j(x)}{t} < 0, \qquad \forall j \in J(x) = \{j \in \{1,\dots,m\} : g_j(x) = 0\};$$

where $T_X(x)$ *is the contingent tangent cone of* X *at* x.

(ii) when X *is convex and* $g_j(j = 1, ..., m)$ *are strictly quasi convex or* g_j, $j = 1, ..., m$ *are convex, the Slater constraint qualification holds:* $X_1 \neq \emptyset$.

5. If the relation (3.5) and (3.12), (3.13), (3.14) hold, then we get

$$\lim_{\epsilon \downarrow 0} \theta_\epsilon = M_P.$$

Indeed, by (3.14), we have

$$\theta_\epsilon \leq M_P + \epsilon.$$

It follows that

$$\limsup_{\epsilon \downarrow 0} \theta_\epsilon \leq M_P. \tag{3.19}$$

On the other hand, by setting $\epsilon' = 2\epsilon$ *in (3.12) and (3.13), we get* $x_{2\epsilon} \in X$ *such that*

$$f(x_{2\epsilon}) \leq \theta_\epsilon + \epsilon$$

and

$$g_j(x_{2\epsilon}) \leq 2\epsilon, \; j = 1, \cdots, m.$$

These combined with (3.5) yield

$$M_P \leq \liminf_{\epsilon \downarrow 0} \theta_\epsilon. \tag{3.20}$$

(3.19) and (3.20) give us $\lim_{\epsilon \downarrow 0} \theta_\epsilon = M_P$.

This further implies that under condition (C_1) or (C_2) and condition (3.5), the ϵ-roots of ϕ_ϵ approach the optimal value M_P of (P) as $\epsilon \downarrow 0$.

4. Approximate Solutions and Nonlinear Lagrangian

In this section, we deal with the relationship between almost approximate solutions of (P) and approximate solutions of a nonlinear penalty problem.

Let $Y \subset R^{m+1}$. A function $p : R^{m+1} \to R$ is called *increasing* if

$$\forall y^1, y^2 \in Y \text{ with } y^1 - y^2 \in R_+^{m+1} \text{ implies } p(y^1) \geq p(y^2).$$

In this section, we will consider increasing and l.s.c. functions p defined on R^{m+1}, which enjoy the following properties:

(A) There exist positive real numbers $a_1, \ldots, a_m$ such that for any $y = (y_0, y_1, ..., y_m)$ belonging to the domain of p with $y_0 \in R_+$, we have

$$p(y) \geq \max\{y_0, a_1 y_1, \ldots, a_m y_m\}.$$

(B) For any $y_0 \in R_+$,

$$p(y_0, 0, \ldots, 0) = y_0.$$

Let

$$F(x, d) = (f(x), d_1 g_1(x), \ldots, d_m g_m(x)), \quad \forall x \in X,\ d = (d_1, ..., d_m) \in R_+^m.$$

We call $L(x, d) = p(F(x, d))$ a *nonlinear Lagrangian* corresponding to p.

PROPOSITION 4.1 *[5]. $L(., d)$ is l.s.c. for any $d \in R_+^m$.*

PROPOSITION 4.2 *[13]. $L(x, d) = f(x), \forall x \in X_0,\ d \in R_+^m$.*

Consider the following nonlinear penalty problem (Q_d):

$$\inf_{x \in X} L(x, d),$$

where $L(x, d) = p(f(x), d_1 g_1(x), ..., d_m g_m(x))$, p is the increasing function defined as above and $d = (d_1, ..., d_m) \in R_+^m$.

PROPOSITION 4.3 *For any $\epsilon > 0$, there exists $d(\epsilon) \in R_+^m$ such that whenever $d - d(\epsilon) \in R_+^m$, every ϵ-solution of (Q_d) is an almost ϵ-solution of (P).*

Proof. For any $d \in R_+^m$, there exists $x_d^\epsilon \in X$, which is an ϵ-solution of $L(x, d)$ on X, such that

$$p(f(x_d^\epsilon), d_1 g_1(x_d^\epsilon), ..., d_m g_m(x_d^\epsilon)) = L(x_d^\epsilon, d) \leq L(x, d) + \epsilon, \qquad \forall x \in X. \tag{4.21}$$

By property (A) of p, $\exists a_j > 0$, $j = 1, ..., m$ such that

$$\max\{f(x), a_1 d_1 g_1(x), ..., a_m d_m g_m(x)\} \\ \leq p(f(x), d_1 g_1(x), ..., d_m g_m(x)), \forall x \in X. \tag{4.22}$$

(4.21) and (4.22) jointly yield

$$\max\{f(x_d^\epsilon), a_1 d_1 g_1(x_d^\epsilon), ..., a_m d_m g_m(x_d^\epsilon)\} \leq \inf_{x \in X} L(x, d) + \epsilon. \tag{4.23}$$

Thus, we have

$$\max\{f(x_d^\epsilon), a_1 d_1 g_1(x_d^\epsilon), ..., a_m d_m g_m(x_d^\epsilon)\} \leq M_P + \epsilon$$

since $\inf_{x \in X} L(x, d) \leq M_P$. So we get

$$f(x_d^\epsilon) \leq M_P + \epsilon, \tag{4.24}$$

$$\max_{1\le j\le m}\{a_j d_j g_j(x_d^\epsilon)\} \le M_P + \epsilon. \tag{4.25}$$

From (4.25), we get

$$[\min_{1\le j\le m} a_j]\cdot[\min_{1\le j\le m} d_j^\epsilon]\cdot g(x_d^\epsilon) \le M_P + \epsilon.$$

Therefore, we can choose $d(\epsilon) = (d_1^\epsilon, ..., d_m^\epsilon) \in R_+^m$ such that

$$\frac{M_P + \epsilon}{[\min_{1\le j\le m} a_j]\cdot[\min_{1\le j\le m} d_j^\epsilon]} < \epsilon,$$

namely, $x_d^\epsilon \in X(\epsilon)$. This combined with (4.24) yields that x_d^ϵ is an almost ϵ-solution of (P). □

REMARK 4.1 *Let the set-valued map $X(\epsilon)$ be u.s.c. at 0 and f uniformly continuous on a neighborhood U of X_0 and $\epsilon_k \to 0^+$. Let $d^k = (d_1^k, ..., d_m^k) \to +\infty$ (i.e., $d_j^k \to +\infty$ as $k \to +\infty$, $j = 1, ..., m$) be such that $x_{d^k}^{\epsilon_k} \in X(\epsilon_k)$. By the u.s.c. of $X(\epsilon)$ at 0, there exists $x^k \in X_0$ such that*

$$\|x_{d^k}^{\epsilon_k} - x^k\| \to 0 \text{ as } k \to +\infty. \tag{4.26}$$

It follows from (4.24) and (4.26) and the uniform continuity of f on U that

$$f(x_{d^k}^{\epsilon_k}) \to M_P.$$

PROPOSITION 4.4 *Let $\epsilon > 0$. Then there exists $d(\epsilon) \in R_+^m$ such that whenever $d - d(\epsilon) \in R_+^m$, every ϵ-regular solution of (Q_d) is an almost regular ϵ-solution of (P).*

Proof. Let $d \in R_+^m$. Note that $L(., d)$ is still a l.s.c. function (by Proposition 4.1) such that $\inf_{x\in X} L(x, d) \ge \inf_{x\in X} f(x) > 0$ according to our assumption. By Proposition 2.1, $\forall\epsilon > 0$, $\exists x_d^\epsilon \in X$ such that

$$L(x_d^\epsilon, d) \le \inf_{x\in X} L(x, d) + \epsilon, \tag{4.27}$$

and

$$L(x_d^\epsilon, d) \le L(x, d) + \epsilon\|x - x_d^\epsilon\|, \quad \forall x \in X. \tag{4.28}$$

It follows from (4.27) and Proposition 4.1 that $\exists d(\epsilon) \in R_+^m$ such that whenever $d - d(\epsilon) \in R_+^m$, x_d^ϵ is an almost ϵ-solution of (P). (4.28) and Proposition 4.2 as well as property (A) of the function p jointly yield that

$$f(x_d^\epsilon) \le L(x_d^\epsilon, d) \le f(x) + \epsilon\|x - x_d^\epsilon\|, \quad \forall x \in X_0.$$

So we conclude that when $d - d(\epsilon) \in R^m_+$, x^ϵ_d is an almost regular solution of (P). □

For illustration, we now apply Proposition 4.4 by taking

$$L^r(x, d) = [f^r(x) + \sum_{j=1}^m d^r_j g^{+r}_j(x)]^{1/r}, \quad \forall x \in X, d = (d_1, ..., d_m) \in R^m_+,$$

where $r > 1$, to derive a so-called generalized KKT condition up to precision ϵ (see, e.g., [1]) under the assumption that f, g_j are locally Lipschitz.

PROPOSITION 4.5 *Let $\epsilon > 0$. Then there exists an almost regular approximate solution x_ϵ for (P) and real numbers $\mu_j(\epsilon) \geq 0$, $j = 1, ..., m$ such that*

(i) $\mu_j(\epsilon) = 0$ if $g_j(x_\epsilon) \leq 0$;
(ii) $\mu_j(\epsilon) > 0$ if $j \in J(\epsilon) = \{j : 0 < g_j(x_\epsilon) \leq \epsilon\}$;
(iii) $0 \in \partial f(x_\epsilon) + \sum_{j \in J(\epsilon)} \mu_j(\epsilon)\partial g_j(x_\epsilon) + \epsilon B^* + N^C_X(x_\epsilon)$, where $\partial f(x)$ denotes the Clarke generalized subdifferetial of f at x, B^* is the unit ball of R^n and $N^C_X(x)$ stands for the Clarke normal cone of X at $x \in X$.

Proof. Let $L^r(x, d)$ be selected as above. According to Proposition 4.2, $\exists d(\epsilon) = (d^\epsilon_1, ..., d^\epsilon_m) \in R^m_+$ and $x_\epsilon \in X$ such that x_ϵ is a regular ϵ-solution of $(Q_{d(\epsilon)})$ and an almost regular ϵ-solution of (P). The former tells us that x_ϵ solves $\min_{x \in X}[L^r(x, d) + \epsilon\|x - x_\epsilon\|]$. Note that $L^r(x, d)$ is still locally Lipschitz. Applying the corollary of Proposition 2.4.3 in [8], we get

$$\begin{aligned} 0 &\in \partial_x L^r(x_\epsilon, d) + \epsilon B^* + N^C_X(x_\epsilon) \\ &\subseteq s[f^{r-1}(x_\epsilon)\partial f(x_\epsilon) + \sum_{j \in J(\epsilon)} (d^\epsilon_j)^k g^{r-1}_j(x_\epsilon)\partial g_j(x_\epsilon)] + \epsilon B^* + N^C_X(x_\epsilon), \end{aligned}$$

where $s = [f^r(x_\epsilon) + \sum_{j=1}^m (d^\epsilon_j)^r g^{+r}_j(x_\epsilon)]^{1/r-1}$.

Let

$$\mu_j(\epsilon) = \frac{(d^\epsilon_j)^r \cdot g^{r-1}_j(x_\epsilon)}{s f^{r-1}(x_\epsilon)}, \quad \text{if } j \in J(\epsilon)$$

and

$$\mu_j(\epsilon) = 0, \quad \text{if } g_j(x_\epsilon) \leq 0.$$

Then

$$0 \in \partial f(x_\epsilon) + \sum_{j \in J(\epsilon)} \mu_j(\epsilon)\partial g_j(x_\epsilon) + \epsilon' B^* + N^C_X(x_\epsilon), \tag{4.29}$$

where $\epsilon' = \dfrac{\epsilon}{sf^{r-1}(x_\epsilon)}$.

Note that

$$sf^{r-1}(x_\epsilon) \geq [f^r(x_\epsilon)]^{1/r-1} \cdot f^{r-1}(x_\epsilon) = 1.$$

Thus, $\epsilon' < \epsilon$. It follows from (4.29) that (iii) holds. □

REMARK 4.2 *1. Taking different forms of nonlinear Lagrangian function $L(x, d)$ may yield different approximate optimality conditions for (P).*

2. If f, g_j are continuously differentiable, (iii) becomes

$$-[\bigtriangledown f(x_\epsilon) + \sum_{j\in J(\epsilon)} \mu_j(\epsilon) \bigtriangledown g_j(x_\epsilon)] \in \epsilon B^* + N_X^C(x_\epsilon).$$

Furthermore, if $X = R^n$, then (iii) becomes

$$\| \bigtriangledown f(x_\epsilon) + \sum_{j\in J(\epsilon)} \mu_j(\epsilon) \bigtriangledown g_j(x_\epsilon)\| \leq \epsilon.$$

3. It follows from Proposition 2.1 that there exists $x^* \in X_0$ which solves

$$\begin{array}{ll} \text{min} & f(x) + \epsilon\|x - x^*\| \\ \text{s.t.} & x \in X, \\ & g_j(x) \leq 0,\ j = 1, ..., m. \end{array}$$

Further, assume that f, g_j are locally Lipschitz, applying Theorem 6.1.1 in [1], we obtain $\lambda \geq 0$, $\mu_j \geq 0$, $j = 1, ..., m$, not all zero, such that

$$\mu_j = 0, \quad \text{if } g_j(x^*) > 0$$

and

$$0 \in \lambda\partial f(x^*) + \sum_{j=1}^{m} \mu_j \partial g_j(x^*) + \lambda\epsilon B^* + N_X^C(x^*),$$

which is a Fritz-John type condition up to precision ϵ. To obtain the KKT type condition up to precision ϵ, i.e., $\lambda \neq 0$, one must impose some constraint qualification (see Remark 5.5 in [9]).

5. Conclusions

In this paper, we investigated approximate optimal solutions of a constrained mathematical programming problems via nonlinear Lagrangian functions. We discussed the relationship between approximate optimal values and approximate roots of a nonlinear function. The almost approximate optimal solutions and approximate generalized KKT type optimality conditions for a constrained optimization problem were studied by means of nonlinear Lagrangian functions.

References

1. Clarke, F.H. (1983), Optimization and Nonsmooth Analysis, John-Wiley & Sons, New York.
2. Ekeland, I. (1974), On variational principle, *J. Math. Anal. Appl.* 47: 324–353.
3. Goh, C.J. and Yang, X.Q. (1997), A sufficient and necessary condition for nonconvex constrained optimization, Applied Mathematics Letters 10: 9–12.
4. Huang, X.X. and Yang, X.Q., Duality of multiobjective optimization via nonlinear Lagrangian functions (submitted for publication).
5. Huang, X.X. and Yang, X.Q., Nonlinear Lagrangian for multiobjective optimization problems and applications to duality and exact penalization (submitted for publication).
6. Li, D. and Sun, X.L. (1999), Value-estimation function method for constrained global optimization, *JOTA*, 102: 385–409.
7. Liu, J.C. (1991), ϵ-duality theorem of nondifferentiable nonconvex multiobjective programming, *JOTA* 69, 153–167.
8. Liu, J. C. (1996), ϵ-Pareto optimality for nondifferentiable programming via penalty function, *J. Math. Anal. Appl.*, 198: 248–261.
9. Loridan, P. (1982), Necessary conditions for ϵ-optimality, Mathematical Programming Study, 19: 140–152.
10. Loridan, P. (1984), ϵ-solution in vector minimization problems, JOTA 43, 265–267.
11. Loridan, P. and Morgan, J. (1983), Penalty functions in ϵ-programming and ϵ-minimax problems, *Math. Programming* 26: 213–231.
12. Rockafellar, R.T. and Wets, R.J-B. (1998), Variational Analysis, Springer, Berlin.
13. Rubinov, A.M., Glover, B.M. and Yang, X.Q. (1999), Modified Lagrangian and penalty functions in continuous optimization, *Optimization* 46: 327–351.
14. Rubinov, A.M., Glover, B.M. and Yang, X.Q. (1999) Decreasing functions with applications to penalization, *SIAM J. Optimization* 10, (1): pp. 289–313.
15. Rubinov, A.M., Yang, X.Q. and Glover, B.M., Nonlinear constrained optimization methods: a review, in *Progress in Optimization: Contributions from Australia*, edited by Yang X.Q. et al.
16. Strodiot, J.J., Nguyen, V.H. and Heukemes, N. (1983), ϵ-optimal solutions in nondifferentiable convex programming and some related questions, *Math. Programming*, 25: 307–328.
17. Yang, X.Q. and Li, D. (2000), Successive global optimization method for constrained global optimization, *Journal of Global Optimization*, 16: 355–369.
18. Yang, X.Q. and Huang, X.X., A nonlinear Lagrangian approach to constrained optimization problems, *SIAM J. Optimization* (accepted).
19. Yokoyama, K. (1992), ϵ-Optimality criteria for convex programming problems via exact penalty functions, *Math. Programming* 56: 233–243.
20. Yokoyama, K. (1994), ϵ-Optimality criteria for vector minimization problems via exact penalty functions, *J. Math. Anal. Appl.* 187, 296–305.