Journal of Global Optimization **16:** 355–369, 2000.

Successive Optimization Method via Parametric Monotone Composition Formulation*

X.Q. YANG[1] and D. LI[2]
[1]*Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong*
[2]*Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong (e-mail: dli@se.cuhk.edu.hk)*

(Received 16 February 1999; accepted in revised form 6 December 1999)

Abstract. In this paper a successive optimization method for solving inequality constrained optimization problems is introduced via a parametric monotone composition reformulation. The global optimal value of the original constrained optimization problem is shown to be the least root of the optimal value function of an auxiliary parametric optimization problem, thus can be found via a bisection method. The parametric optimization subproblem is formulated in such a way that it is a one-parameter problem and its value function is a monotone composition function with respect to the original objective function and the constraints. Various forms can be taken in the parametric optimization problem in accordance with a special structure of the original optimization problem, and in some cases, the parametric optimization problems are convex composite ones. Finally, the parametric monotone composite reformulation is applied to study local optimality.

Key words: Global optimization; Parametric monotone composition; Convexification; Optimality condition; Least root problem

1. Introduction

Successful applications of optimization [1, 3, 5, 16, 17] can be found in various areas, such as finance, engineering and operations research. During the last two decades, the study of global optimization has attracted a lot of attention from optimization researchers. Solution methods for various nonconvex or discrete optimization problems have been developed, see [6, 12, 13, 18]. In particular, convexification methods in nonconvex optimization are of considerable interest to optimization researchers, as there is no standard algorithm for nonconvex optimization problems. Various convexification methods have been studied in [1, 11]. Exponential transformation and pth power transformation are two commonly used convexification techniques.

Consider the following optimization problem P

* This research was partially supported by the Australian Research Council and the Research Grants Council of Hong Kong (grant no. CUHK 358/96P).

$$\begin{aligned} &\inf \quad f_0(x) \\ &\text{subject to} \quad f_j(x) \leq 0 \quad j = 1, \ldots, m\,, \\ &\qquad x \in X\,, \end{aligned}$$

where X is a subset of $\mathbb{R}^n$, $f_i : \mathbb{R}^n \to \mathbb{R}$ ($i = 0, 1, \ldots, m$) are real-valued functions.

If the perturbation function for P is not convex, the primal-dual methods [16] may fail to solve the problem P. Recently, a pth power transformation was developed in [14] to convexify the perturbation function and thus to achieve a property of zero duality gap for a class of nonconvex optimization problems. In [19, 22], a nonlinear Lagrangian dual formulation with a minimax or pth power structure for nonconvex continuous or discrete inequality constrained optimization problems is introduced and zero duality gap is shown to exist under very mild condition. In particular, by virtue of the perturbation function exact penalty results via nonlinear penalty functions are provided in [19] with smaller penalty parameters than the ones that are required by the classical penalty functions. It is worth noting that these zero duality results are motivated by a recent sufficient and necessary condition for a nonconvex optimization problem obtained in [7]. It is also worth noting that in [4] a problem of minimizing a concave quadratic function subject to finitely many convex quadratic constraints is reconstructed as an equivalent minimax convex problem.

The aim of this paper is to show that finding the global optimal value of constrained inequality optimization problems is equivalent to the problem of finding the least root of a nonnegative monotone decreasing function. The least root problem can be solved via, e.g., a bisection method and, however, the value of the nonnegative monotone decreasing function involves a global optimization problem. More precisely, a class of parametric monotone composition formulations for (nonconvex) optimization problems is introduced. The motivation is to refine the sufficient and necessary condition for solving inequality constrained optimization problems in [7].

It is worth noting that the local search scheme may be needed in some global optimization method (see [20]). Thus, the transformed problem in [7] with a minimax structure may have some disadvantages when the local search is applied. The parametric optimization problem to be introduced in this paper is formulated in such a way that it is a one-parameter problem and its value function is a monotone composition function with respect to the original objective function and the constraints. Various monotone composition forms can be taken in the parametric optimization problem formulation. In particular, two elementary transformation functions (exponential and pth power) used in [11] are examples of outer functions for monotone composition problems. The proposed method is a two-level scheme. In the lower level of each iteration, an auxiliary parametric optimization problem with simple constraints or without constraint is solved, while in the upper level the parameter is adjusted, via a bisection method, such that the least root of the optimal value function of the parametric optimization problem is found.

The outline of the paper is as follows. In Section 2, a parametric monotone composition approach for solving inequality constrained optimization problems is introduced. In Section 3, the inequality constrained global optimization problem is shown to be equivalent to the least root problem of the optimal value function of an auxiliary parametric optimization problem. A two-level solution scheme is proposed. In Section 4, an algorithm is designed using the proposed two-level solution scheme and an example is given to illustrate the proposed solution scheme. Finally, local optimality via the parametric monotone composition approach is discussed in Section 5.

2. Unconstrained Monotone Composition Formulation

Consider the following optimization problem P that could be nonconvex

$$\begin{aligned} \inf \quad & f_0(x) \\ \text{subject to} \quad & f_j(x) \leqslant 0 \quad j = 1, \ldots, m\,, \\ & x \in X\,, \end{aligned}$$

where X is a subset of $\mathbb{R}^n$, $f_i : \mathbb{R}^n \to \mathbb{R}$ $(i = 0, 1, \ldots, m)$ are real-valued functions. The feasible set of P is

$$X_0 = \{x \in X\colon\ f_i(x) \leqslant 0, \quad i = 1, \ldots, m\}\,.$$

Our study in this paper is restricted to the class of optimization problems P that satisfy the following assumptions.

(A1) f_0 is uniformly continuous, and each f_i is continuous on an open set containing X, $i = 1, \ldots, m$;

(A2) $\inf_{x \in X} f_0(x) > 0$;

(A3) X is connected.

Uniform continuity of f_0 in Assumption (A1) is an essential assumption in our approach, although it is a very general assumption made in optimization problems. Assumption (A2) states that f_0 is bounded from below on X. The connectedness of X in Assumption (A3) will be used in the proof of Lemma 3.2.

DEFINITION 2.1. *Let the function* $\phi : \mathbb{R}^{1+m} \to \mathbb{R}$ *satisfy the following properties, for* $y = (y_0, y_1, \ldots, y_m)$*:*

(A) $\phi(y) \geqslant \max_{0 \leqslant i \leqslant m} y_i$, $\forall y \in \mathbb{R}^{1+m}$;

(B) $\phi(y) = \max\{0, y_0\}$, $\forall y_0 \in \mathbb{R}$, $(y_1, \ldots, y_m) \in \mathbb{R}^m_-$;

(C) ϕ *is a nondecreasing function with respect to its first component* y_0.

It is clear from (B) that $\phi(0, 0, \ldots, 0) = \max\{0, 0\} = 0$.

EXAMPLE 2.1. The following functions satisfy properties (A), (B) and (C):

$$\phi_\infty(y) = \max\{0, y_0, y_1, \ldots, y_m\};$$

$$\phi_p(y) = \left(\sum_{i=0}^{m} [\max\{0, y_i\}]^p\right)^{\frac{1}{p}}, \quad 0 < p < +\infty;$$

$$\phi'_p(y) = \max\{0, y_0\} + \left(\sum_{i=1}^{m} [\max\{0, q_i(y_i)\}]^p\right)^{\frac{1}{p}}, \quad 0 < p < +\infty,$$

where $q_i : \mathbb{R} \to \mathbb{R}$ can be any continuous function that satisfies $q_i(z) \leqslant 0$, if $z < 0$ and $q_i(z) \geqslant z$, if $z \geqslant 0$. For example, $q_i(z) = \exp(z) - 1$. □

Let

$$F(x, \theta_0) = \left(\frac{f_0(x)}{\theta_0} - 1, f_1(x), \ldots, f_m(x)\right), \quad x \in X, \theta > 0.$$

DEFINITION 2.2. *Let ϕ be a function that satisfies properties (A), (B) and (C) in Definition 2.1 and $\theta_0 > 0$. Define an auxiliary function $\Phi : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ as*

$$\Phi(x, \theta_0) = \phi(F(x, \theta_0)).$$

An auxiliary problem $P_\phi(\theta_0)$ is defined as follows

$$\inf_{x \in X} \Phi(x, \theta_0).$$

Constraint $x \in X$ often represents simple constraints such as constraints for lower and upper bounds. We expect the constraint structure in $P_\phi(\theta_0)$ to be much simpler than the constraint structure in P in general situations. This form of the auxiliary parametric optimization problem can be considered as a monotone composition formulation since ϕ is a monotone function with respect to its first component. When $X = \mathbb{R}^n$, $P_\phi(\theta_0)$ becomes an unconstrained optimization problem.

The problem $P_\phi(\theta_0)$ can be also rewritten as the following unconstrained optimization problem,

$$\inf_{x \in \mathbb{R}^n} \phi(F(x, \theta_0)) + \delta_X(x),$$

where δ_X is an indicator function of X: $\delta_X(x) = 0$, if $x \in X$ and $\delta_X(x) = +\infty$, if $x \not\in X$. Furthermore, if ϕ is a convex function and X is a convex set, then the problem $P_\phi(\theta_0)$ is a convex composite optimization problem of which global second-order sufficient optimality conditions have been studied in [21].

In particular, if $\phi = \phi_\infty$, problem $P_\phi(\theta_0)$ becomes

$$\inf_{x \in X} \max\left\{0, \frac{f_0(x)}{\theta_0} - 1, f_1(x), \ldots, f_m(x)\right\}.$$

This can be considered as a generalized minimax formulation for solving P. If $\phi = \phi_2$, then the auxiliary problem $P_\phi(\theta_0)$ is given by

$$\inf_{x \in X} \left\{ \left[\max\left\{0, \frac{f_0(x)}{\theta_0} - 1 \right\} \right]^2 + \sum_{i=1}^{m} [\max\{0, f_i(x)\}]^2 \right\}^{1/2} .$$

The above problem can be further reduced to the following equivalent form (without the square root),

$$\inf_{x \in X} \left[\max\left\{0, \frac{f_0(x)}{\theta_0} - 1 \right\} \right]^2 + \sum_{i=1}^{m} [\max\{0, f_i(x)\}]^2 .$$

Note that the above equivalent formulation of $P_\phi(\theta_0)$ yields a differentiable optimization problem if all the functions f_j's are differentiable.

It is worth noting that the function defined by

$$\phi(y) = \max\{y_0, y_1, \ldots, y_m\} ,$$

does not satisfy property (B) and that this nonsmooth function is used in [7] to formulate another type of the auxiliary parametric optimization problems for problem P. An alternative approach using a minimax formulation was given in [10]. If the global optimal solution x^* is known, the problem P is equivalent to the following unconstrained optimization problem (see [10])

$$\min_{x \in X} \max\{f_0(x) - f_0(x^*), f_1(x), \ldots, f_m(x)\} .$$

The following lemma of properties of the auxiliary function $\Phi(x, \theta_0)$ will be often used in sequel.

LEMMA 2.1. *Let $\theta_0 > 0$ and $x \in X$. Then $\Phi(x, \theta_0) \geq 0$. Furthermore if $\Phi(x, \theta_0) = 0$, then $x \in X_0$.*

Proof. We consider two cases:

Case 1: x is infeasible. Then from property (A) for some $1 \leq i \leq m$

$$\phi\left(\frac{f_0(x)}{\theta_0} - 1, f_1(x), \ldots, f_m(x)\right) \geq f_i(x) > 0 . \tag{1}$$

Case 2: x is feasible. Then from property (B)

$$\phi\left(\frac{f_0(x)}{\theta_0} - 1, f_1(x), \ldots, f_m(x)\right) = \max\left\{0, \frac{f_0(x)}{\theta_0} - 1 \right\} \geq 0 .$$

Assume that $\Phi(x, \theta_0) = 0$. If x is infeasible, then from (1), $\Phi(x, \theta_0) > 0$, which is a contradiction. □

3. Global Optimality and Least Root Problem

In this section, a successive solution scheme via parametric monotone composition formulation is developed for finding a global minimum of optimization problems that could be nonconvex. More specifically, a two-level iterative scheme will be proposed. In the lower level of each iteration, an auxiliary optimization problem with a fixed parameter is solved, while in the upper level the parameter is adjusted through finding the least root of the optimal value function of the parametric optimization problem. It is worth noting that although the global optimal solution for P is assumed in the following, the algorithm (see Section 4) does not need to know that as a priori.

DEFINITION 3.1. *Let ϕ be a function that satisfies properties (A), (B) and (C) and $\theta_0 > 0$. Define $\varphi(\theta_0)$ to be the global optimal value of $P_\phi(\theta_0)$, i.e.,*

$$\varphi(\theta_0) = \inf_{x \in X} \phi(F(x, \theta_0)) .$$

In the following lemmas, let ϕ be a function satisfying (A), (B) and (C). The following result provides a sufficient and necessary optimality condition for P which refines the one in [7].

LEMMA 3.1. *Let x^* solve the problem P and $\theta_0^* = f_0(x^*)$. Then x^0 solves P if and only if x^0 solves the problem $P_\phi(\theta_0^*)$.*

Proof. From Assumption (A2), $\theta_0^* > 0$. It is clear that $x^* \in X_0$ and

$$\Phi(x^*, \theta_0^*) = \phi\left(\frac{f_0(x^*)}{\theta_0^*} - 1, f_1(x^*), \ldots, f_m(x^*)\right) = \max\left\{0, \frac{f_0(x^*)}{\theta_0^*} - 1\right\} = 0 .$$

From Lemma 2.1, $\Phi(x, \theta_0^*) \geq 0$, $\forall x \in X$. Thus the optimal value of the problem $P_\phi(\theta_0^*)$ is 0.

If x^0 solves P, then from the above, $\Phi(x^0, \theta_0^*) = 0$, thus x^0 solves $P_\phi(\theta_0^*)$.

If x^0 does not solve P, then there are two cases:

Case 1. x^0 is infeasible, then from property (A) for some $1 \leq i \leq m$

$$\phi\left(\frac{f_0(x^0)}{\theta_0^*} - 1, f_1(x^0), \ldots, f_m(x^0)\right) \geq f_i(x^0) > 0 .$$

Case 2. x^0 is feasible, then $f_0(x^0) > f_0(x^*) = \theta_0^*$. Thus

$$\phi\left(\frac{f_0(x^0)}{\theta_0^*} - 1, f_1(x^0), \ldots, f_m(x^0)\right) \geq \frac{f_0(x^0)}{\theta_0^*} - 1 > 0 .$$

Then x^0 does not solve $P_\phi(\theta_0^*)$. □

LEMMA 3.2. *Let x^* solve the problem P. If $0 < \theta_0 < f_0(x^*)$, then $\varphi(\theta_0) > 0$.*

Proof. Let $\bar{X}_0 = X \backslash X_0$. Assume that $0 < \theta_0 < f_0(x^*)$. It is clear that

$$\begin{aligned}\varphi(\theta_0) &= \inf_{x \in X} \phi(F(x, \theta_0)) \\ &= \min\{\inf_{x \in X_0} \phi(F(x, \theta_0)), \inf_{x \in \bar{X}_0} \phi(F(x, \theta_0))\} .\end{aligned}$$

From property (B), it is easy to see that

$$\inf_{x\in X_0} \phi(F(x,\theta_0)) = \inf_{x\in X_0} \max\left\{0, \frac{f_0(x)}{\theta_0} - 1\right\} \geq \frac{f_0(x^*)}{\theta_0} - 1 > 0\,. \tag{2}$$

Let us consider the term $\inf_{x\in \bar{X}_0} \phi(F(x,\theta_0))$. For any $\delta > 0$, the set $\bar{X}_0$ can be decomposed into

$$\bar{X}_0 = \bar{X}_0^{>\delta} \cup \bar{X}_0^{<\delta}\,,$$

where $\bar{X}_0^{>\delta} = \{x \in \bar{X}_0 : \exists 1 \leq i \leq m,\ f_i(x) > \delta\}$ and $\bar{X}_0^{<\delta} = \{x \in \bar{X}_0 : \forall 1 \leq i \leq m,\ f_i(x) \leq \delta\}$. Then

$$\inf_{x\in \bar{X}_0} \phi(F(x,\theta_0)) = \min\{\inf_{x\in \bar{X}_0^{>\delta}} \phi(F(x,\theta_0)), \inf_{x\in \bar{X}_0^{<\delta}} \phi(F(x,\theta_0))\}\,.$$

We have from property (A)

$$\phi(F(x,\theta_0)) \geq f_{i(x)}(x) > \delta\,, \quad \forall x \in \bar{X}_0^{>\delta} \text{ and some } i(x)\,.$$

Thus,

$$\inf_{x\in \bar{X}_0^{>\delta}} \phi(F(x,\theta_0)) \geq \delta > 0\,. \tag{3}$$

Note that f_0 is uniformly continuous and X is connected. For $\frac{f_0(x^*)-\theta_0}{2} > 0$, choose a small $\delta > 0$ such that for any $x \in \bar{X}_0^{<\delta}$, there exists $y \in X_0$ such that

$$|f_0(x) - f_0(y)| \leq \frac{f_0(x^*) - \theta_0}{2}\,.$$

Then

$$f_0(x) \geq f_0(y) - \frac{f_0(x^*)-\theta_0}{2} \geq f_0(x^*) - \frac{f_0(x^*)-\theta_0}{2} = \frac{f_0(x^*)+\theta_0}{2}\,.$$

We further have

$$\frac{f_0(x)}{\theta_0} - 1 \geq \frac{f_0(x^*)-\theta_0}{2\theta_0}\,, \quad \forall x \in \bar{X}_0^{<\delta}\,,$$

i.e.,

$$\inf_{x\in \bar{X}_1^{<\delta}} \phi(F(x,\theta_0)) \geq \frac{f_0(x^*)-\theta_0}{2\theta_0} > 0\,. \tag{4}$$

Summarizing (2), (3), (4), we have that $\varphi(\theta_0) > 0$. □

LEMMA 3.3. *Let x^* solve the problem P. If $\theta_0 \geq f_0(x^*)$, then $\varphi(\theta_0) = 0$.*

Proof. Let $\theta_0 \geq f_0(x^*)$. If x is feasible, then from property (B),

$$\phi\left(\frac{f_0(x)}{\theta_0} - 1, f_1(x), \ldots, f_m(x)\right) = \max\left\{0, \frac{f_0(x)}{\theta_0} - 1\right\} \geq 0\,.$$

If x is infeasible, then there exists $1 \leqslant i \leqslant m$ such that $f_i(x) > 0$, and so from property (A)

$$\phi\left(\frac{f_0(x)}{\theta_0} - 1,\ f_1(x), \ldots, f_m(x)\right) \geqslant f_i(x) > 0\,.$$

Thus, $\Phi(x, \theta_0) \geqslant 0$. Note that

$$\phi\left(\frac{f_0(x^*)}{\theta_0} - 1,\ f_1(x^*), \ldots, f_m(x^*)\right) = \max\left\{0, \frac{f_0(x^*)}{\theta_0} - 1\right\} = 0\,.$$

Then $\varphi(\theta_0) = 0$. □

LEMMA 3.4. *$\varphi(\theta_0)$ is a continuous, nonincreasing and nonnegative function of θ_0 ($\theta_0 > 0$).*

Proof. The continuity of $\varphi(\theta_0)$ follows from Theorem 4.2.1 of [2]. The nonincreasing property of $\varphi(\theta_0)$ is assured by property (C). Lemmas 3.2 and 3.3 together show that $\varphi(\theta_0) \geqslant 0$, $\forall \theta_0 > 0$. □

The main result of the paper is given below.

THEOREM 3.1. *Let ϕ be a function satisfying (A), (B) and (C). θ_0 is the global optimal value of P if and only if θ_0 is the solution of the following least root problem*

$$\theta_0 = \min\{\theta \mid \varphi(\theta) = 0\}\,. \tag{5}$$

Proof. The result follows from Lemmas 3.2 and 3.3. □

REMARK 3.1. Although it is intuitively straightforward to find the least root of a nonincreasing function in problem (5), finding the global optimum of the auxiliary problem $P_\phi(\theta_0)$ is, in general, not an easy task, if the objective function of $P_\phi(\theta_0)$ is not convex. However, the outer function ϕ can have the form of two elementary convexification transformation functions (exponential and pth power) that were used in [11] for convexification. The subproblem $P_\phi(\theta_0)$ may be then easily solved for some classes of nonconvex optimization problems.

The proposed method can be considered as a two-level scheme. The lower level is to solve, for a given parameter, an optimization problem with simple constraints or without constraint. The upper level is to check if the parameter θ_0 is the least root of the equation $\varphi(\theta_0) = 0$.

COROLLARY 3.1. *θ_0 is the global optimal value of P if and only if θ_0 is the least root of the equation*

$$\varphi_\infty(\theta_0) = 0\,,$$

where $\varphi_\infty(\theta_0)$ is the optimal value of the auxiliary problem $P_\phi^\infty(\theta_0)$:

$$\inf_{x\in X} \max\left\{0, \frac{f_i(x)}{\theta_0} - 1,\ f_1(x), \dots, f_m(x)\right\}.$$

Proof. The result follows from Theorem 3.1 by letting $\phi(y) = \phi_\infty(y)$. □

COROLLARY 3.2. *θ_0 is the global optimal value of P if and only if θ_0 is the least root of the equation*

$$\bar{\varphi}_p(\theta_0) = 0\,, \tag{6}$$

where $\bar{\varphi}_p(\theta_0)$ is the optimal value of the subproblem $\bar{P}^p_\phi(\theta_0)$:

$$\inf_{x\in X} \left[\max\left\{0, \frac{f_0(x)}{\theta_0} - 1\right\}\right]^p + \sum_{i=0}^{m} [\max\{0,\ f_i(x)\}]^p\,. \tag{7}$$

Proof. From Theorem 3.1, $\theta_0 = f_0(x^*)$ if and only if θ_0 is the least root of the equation

$$\varphi_p(\theta_0) = 0\,,$$

where $\varphi_p(\theta_0)$ is the optimal value of the auxiliary problem $P^p_\phi(\theta_0)$:

$$\inf_{x\in X} \left(\left[\max\left\{0, \frac{f_0(x)}{\theta_0} - 1\right\}\right]^p + \sum_{i=0}^{m} [\max\{0,\ f_i(x)\}]^p\right)^{\frac{1}{p}}.$$

This is equivalent to that θ_0 is the least root of the equation

$$(\bar{\varphi}_p(\theta_0))^{\frac{1}{p}} = 0\,,$$

where $\bar{\varphi}_p(\theta_0)$ is the optimal value of the subproblem $\bar{P}^p_\phi(\theta_0)$. It is clear that the above equation is equivalent to (6). □

REMARK 3.2. (i) It is much easier to solve $\bar{P}^p_\phi(\theta_0)$ than $P^p_\phi(\theta_0)$ since there is no pth root in the objective function of $\bar{P}^p_\phi(\theta_0)$.

(ii) Assume that x^* solves P. It is clear that if for $\theta_0 > f_0(x^*)$, $0 < p_1,\ p_2 < +\infty$, the solution x_{p_1} of $\bar{P}^{p_1}_\phi(\theta_0)$ and the solution x_{p_2} of $\bar{P}^{p_2}_\phi(\theta_0)$ are feasible for P, then $\bar{\varphi}_{p_1}(\theta_0) = \bar{\varphi}_{p_2}(\theta_0)$. Thus $\varphi_{p_1}(\theta_0) = \varphi_{p_2}(\theta_0)$.

4. Algorithm and Illustrative Examples

We now design an algorithm using the two-level scheme proposed in the last section and a bisection method. We assume that the global optimal value $\varphi(\theta_0)$ for subproblem $P_\phi(\theta_0)$ is available via some global optimization solver, see e.g. [20]. The parameter ϵ in the following algorithm is the accuracy required in the application.

ALGORITHM

Step 0. Given $\epsilon > 0$. Find θ_1 and θ_2 such that

$$\varphi(\theta_1) > 0 \quad \text{and} \quad \varphi(\theta_2) = 0 .$$

Step 1. Compute $\varphi(\frac{\theta_1 + \theta_2}{2})$ via any global optimization solver.

Step 2. If $\varphi(\frac{\theta_1 + \theta_2}{2}) > 0$, let $\theta_1 \leftarrow \frac{\theta_1 + \theta_2}{2}$ and $\theta_2 = \theta_2$.

If $\varphi(\frac{\theta_1 + \theta_2}{2}) = 0$, let $\theta_1 = \theta_1$ and $\theta_2 \leftarrow \frac{\theta_1 + \theta_2}{2}$.

Step 3. If $|\theta_1 - \theta_2| < \epsilon$, stop. Otherwise return to Step 1.

Consider the following nonconvex optimization problem in [14]

$$\begin{aligned} \min \quad & f_0(x) = 1 - x_1 x_2 , \\ \text{subject to} \quad & f_1(x) = x_1 + 4x_2 - 1 \leqslant 0 , \\ & x_1 \geqslant 0,\ x_2 \geqslant 0 . \end{aligned}$$

The exact optimal solution to the problem is $x_1 = 0.5$ and $x_2 = 0.125$ and the corresponding optimal value is $15/16$.

The direct application of the primal-dual method [16] would fail in the original setting of this example problem as its perturbation function is nonconvex [14] and there is a nonzero duality gap. This example problem was solved in [14] by a pth power Lagrangian method. In the pth power Lagrangian method [14], the value of p needs to be chosen sufficiently large in order to convexify the perturbation function. Although a theoretical lower bound can be derived for p [15, 8], how large is large enough for p could be a thorny issue in computational implementation.

From Corollary 3.2, this problem is equivalent to the least root problem

$$\theta_0 = \min\{\theta \mid \varphi_p(\theta) = 0\}$$

where $\varphi_p(\theta_0) = (\bar{\varphi}_p(\theta_0))^{\frac{1}{p}}$ and $\bar{\varphi}_p(\theta_0)$ is the (global) optimal value of the following auxiliary problem $P^p_\phi(\theta_0)$:

$$\min_{x_1, x_2 \geqslant 0} \left[\max\left\{0, \frac{1 - x_1 x_2}{\theta_0} - 1 \right\} \right]^p + [\max\{0, x_1 + 4x_2 - 1\}]^p .$$

Problem $P^p_\phi(\theta_0)$ is solved using the global optimization solver proposed in [20] in which the local search is performed using the optimization tool box in Matlab [9].

Choose $\epsilon = 10^{-3}$, $p = 2$, $\theta_1 = 0.5$ and $\theta_2 = 2$, the algorithm is terminated in 11 iterations. The optimal value lies in the interval $[0.9365, 0.9380]$ and the approximate optimal solution is $x_1 = 0.4991$ and $x_2 = 0.1253$. The result of the algorithm is plotted in Figure 1.

Let us consider another optimization problem

$$\min_{x \in X} -x + 1 \quad \text{s.t. } x^2 \leqslant 0$$

where $X = \{x \in \mathbb{R} : x \geqslant -\frac{1}{2}\}$.

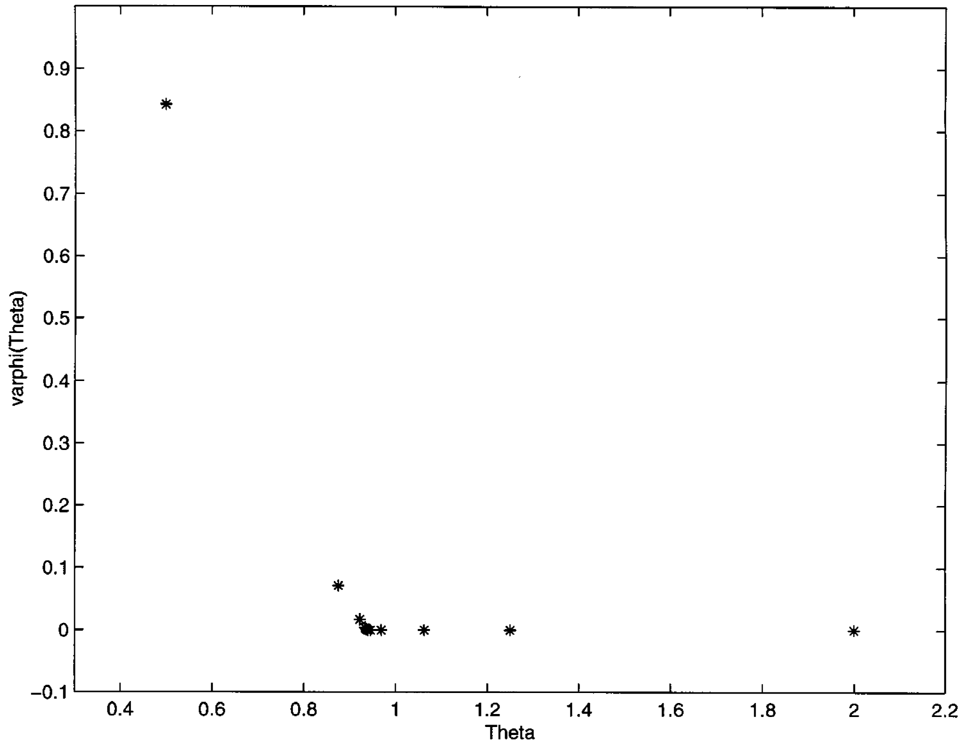

Figure 1. Least root problem via bisection.

For this problem, $x^0 = 0$ is the optimal solution. The primal-dual method [16], however, would still fail as a Lagrange multiplier does not exist.

Furthermore, if the augmented Lagrangian method (p. 414, [16]) is applied to this problem, the dual function for a fixed penalty parameter $c > 0$ is given by

$$\psi(\mu) = \min_{x \in X} (-x + 1 + \mu x^2 + cx^4) .$$

Then the optimal value of the original problem is obtained by solving the following problem

$$\max_{\mu > 0} \psi(\mu) .$$

However, it can be checked that there does not exist an optimal solution to this maximization problem. As a matter of fact, the optimal value of the original problem is attained only when $\mu \to +\infty$. This is due to the fact that the optimal solution $x^0 = 0$ of the original problem is not a regular point [16].

Note that the method proposed in this paper does not require the regularity property. For this example problem, our method results in an auxiliary problem with $p = 1$ as follows:

$$\min_x \left[\max\left\{0, \frac{-x+1}{\theta_0} - 1\right\} \right] + x^2 ,$$

which can be solved easily as it is an unconstrained convex program. The optimal solution $x^0 = 0$ is attained when θ_0 is set to equal to $f_0(x^0) = 1$.

5. Local Optimality

Although the global optimality is the main interest of this paper, we next show that the proposed parametric monotone composition approach is also valid for searching strict local optimality. This concern is of practical significance, since in many situations, only a local optimality of $P_\phi(\theta_0^*)$ can be guaranteed.

THEOREM 5.1. *Consider the problem P. If x^0 is a strict local minimum of P, then x^0 is a strict local minimum of the problem $P_\phi(\theta_0^*)$ with $\Phi(x^0, \theta_0^*) = 0$, where $\theta_0^* = f_0(x^0)$.*

Proof. Assume that x^0 is a strict local minimum of P and $\theta_0^* = f_0(x^0)$. It is clear that

$$\Phi(x^0, \theta_0^*) = \phi\left(\frac{f_0(x^0)}{\theta_0^*} - 1, f_1(x^0), \ldots, f_m(x^0)\right) = \max\left(0, \frac{f_0(x^0)}{\theta_0^*} - 1\right\} = 0 .$$

There is a neighborhood $N_1(x^0)$ of x^0 such that for any $x \in X_0 \cap N_1(x^0)$ and $x \neq x_0$, $f_0(x) > f_0(x^0)$. We will show that for any $x \in X \cap N_1(x^0)$ and $x \neq x_0$, $\Phi(x, \theta_0^*) > 0$.

In fact, if a x satisfying $x \in X \cap N_1(x^0)$ and $x \neq x_0$ is infeasible for P, then from property (A) the following holds for some $1 \leqslant i \leqslant m$

$$\phi\left(\frac{f_0(x)}{\theta_0^*} - 1, f_1(x), \ldots, f_m(x)\right) \geqslant f_i(x) > 0 .$$

If a x satisfying $x \in X \cap N_1(x^0)$ and $x \neq x_0$ is feasible for P, i.e., $x \in X_0 \cap N_1(x^0)$ and $x \neq x_0$, then $f_0(x) > f_0(x^0) = \theta_0^*$. Thus from (A)

$$\phi\left(\frac{f_0(x)}{\theta_0^*} - 1, f_1(x), \ldots, f_m(x)\right) \geqslant \frac{f_0(x)}{\theta_0^*} - 1 > 0 .$$

Thus x^0 is a strict local minimum of the problem $P_\phi(\theta_0^*)$. □

THEOREM 5.2. *Consider the problem P. Let $\theta_0^* > 0$. If x^0 is a strict local minimum of the problem $P_\phi(\theta_0^*)$ and $\Phi(x^0, \theta_0^*) = 0$, then x^0 is a strict local minimum of P.*

Proof. Since $\Phi(x^0, \theta_0^*) = 0$, from Lemma 2.1, x_0 is feasible. It is clear that

$$\Phi(x^0, \theta_0^*) = \max\left\{0, \frac{f_0(x)}{\theta_0^*} - 1\right\} = 0 .$$

Thus $\theta_0^* \geqslant f_0(x^0)$. Assume that there is a neighborhood $N_2(x^0)$ of x^0 such that for any $x \in X \cap N_2(x^0)$ and $x \neq x_0$,

$$\Phi(x, \theta_0^*) > \Phi(x^0, \theta_0^*) = 0 . \tag{8}$$

Note that

$$\Phi(x, \theta_0^*) = \max\left\{0, \frac{f_0(x)}{\theta_0^*} - 1\right\}, \quad \forall x \in X_0 \cap N_2(x^0) .$$

By the strict inequality (8),

$$\Phi(x, \theta_0^*) = \frac{f_0(x)}{\theta_0^*} - 1 > 0 = \Phi(x^0, \theta_0^*), \quad \forall x \in X_0 \cap N_2(x^0), \quad x \neq x^0 .$$

Thus

$$f_0(x) > \theta_0^* \geqslant f_0(x^0), \quad \forall x \in X_0 \cap N_2(x^0), \quad x \neq x^0 .$$

Then x_0 is a strict local minimum of the problem P. □

The condition $\Phi(x^0, \theta_0^*) = 0$ holds if x^0 is feasible and $\theta_0^* = f_0(x^0)$. The following result shows that a local minimum of P is actually a global minimum for $P_\phi(\theta_0^*)$.

THEOREM 5.3. *Consider the problem P. If x^0 is a local minimum of P and $\theta_0^* = f_0(x^0)$, then x^0 is a global minimum of the problem $P_\phi(\theta_0^*)$ with* $\Phi(x^0, \theta_0^*) = 0$.

Proof. Assume that x^0 is a local minimum of P and $\theta_0^* = f_0(x^0)$. It is clear that

$$\Phi(x^0, \theta_0^*) = \phi\left(\frac{f_0(x^0)}{\theta_0^*} - 1, f_1(x^0), \ldots, f_m(x^0)\right) = \max\left\{0, \frac{f_0(x^0)}{\theta_0^*} - 1\right\} = 0 .$$

Then from Lemma 2.1, x^0 is a global minimum of the problem $P_\phi(\theta_0^*)$ with $\Phi(x^0, \theta_0^*) = 0$. □

The following example shows that a local minimum for $P_\phi(\theta_0^*)$ may not be a local minimum for P if the local minimum of $P_\phi(\theta_0^*)$ is not a strict local minimum.

EXAMPLE 5.1. Consider the optimization problem

$$\begin{aligned} \inf \quad & f_0(x) \\ \text{subject to} \quad & x \in X , \end{aligned}$$

where $X = [0, \infty)$ and $f_0(x) = \cos(x)$, if $0 \leqslant x \leqslant 2\pi$ and $x - 2\pi + 1$, if $2\pi \leqslant x$. Let $\theta_0 = 1$. Then $x^0 = 2\pi$ is a local minimum for $P_\phi(1)$:

$$\begin{aligned} \inf \quad & \max\{0, f_0(x) - 1\} , \\ \text{subject to} \quad & x \in X , \end{aligned}$$

where $\max\{0, f_0(x) - 1\} = 0$, if $0 \leqslant x \leqslant 2\pi$ and $x - 2\pi$, if $2\pi \leqslant x$. But $x^0 = 2\pi$ is not a local minimum for the original optimization problem.

6. Conclusions

In this paper, a successive optimization method for solving nonconvex constrained optimization problems was investigated via a parametric monotone composition reformulation and bisection method. The outer functions in this reformulation are of some forms of elementary convexification transformations, thus the transformed subproblem may be convex for some classes of nonconvex optimization problems. The potential advantage of the proposed method is shown by the zero gap property obtained by the nonlinear Lagrangian function in [14, 22] and the exact penalty result with smaller penalty parameters than the ones that are required by the classical penalty function given in [19].

References

1. Avriel, M., Diewert, W.E., Schaible, S. and Zang, I. (1988), *Generalized Concavity*, Plenum Press, New York and London.
2. Bank, B., Guddat, J., Klatte, D., Kummer, B. and Tammer, K. (1983), *Non-linear Parametric Optimization*, Birkhauser Verlag, Basel.
3. Bazaraa, M.S. and Shetty, C.M. (1979), *Nonlinear Programming*, John Wiley & Sons, New York, Singapore.
4. Ben-Tal, A. and Teboulle, M. (1996), Hidden convexity in some nonconvex quadratically constrained quadratic programming, *Mathematical Programming* 72: 51–63.
5. Bertsekas, D.P. (1982), *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York and London.
6. Fongo, F. (1988), *Nonconvex Programming*, Academic Press, Budapest.
7. Goh, C.J. and Yang, X.Q. (1997), A sufficient and necessary condition for nonconvex optimization, *Applied Mathematics Letters* 10: 9–12.
8. Goh, C.J. and Yang, X.Q. (1998), Comment on convexification of a nonconvex frontier, *Journal of Optimization Theory and Applications* 97: 759–768.
9. Grace, A. (1990), *Optimization Toolbox for use with MATLAB user guide*, The Maths Works, Inc., South Natick, MA.
10. Hiriart-Urruty, J.B. and Lemarchal, C. (1993), *Convex analysis and minimization algorithms. I, II*, Springer-Verlag, Berlin.
11. Horst, R. (1984), On the convexification of nonlinear programming problems: An application-oriented survey, *European Journal of Operations Research* 15: 382–392.
12. Horst, R. and Tuy, H. (1991), *Global Optimization – Deterministic Approaches*, Springer-Verlag.
13. Konno, H., Thach, P.T. and Tuy, H. (1997), *Optimization on Low Rank Nonconvex Structures, Kluwer Academic Publishers, Dordrecht/Boston/London*.
14. Li, D. (1995), Zero duality gap for a class of nonconvex optimization problem, *Journal of Optimization Theory and Applications* 85(2): 309–324.
15. Li, D. (1996), Convexification of a noninferior frontier, *Journal of Optimization Theory and Applications* 88: 177–196.

16. Luenberger, D.G. (1989), *Linear and Nonlinear Programming*, second edition, Addison Wesley, Reading, MA.
17. Mangasarian, O.L. (1969), *Nonlinear Programming*, McGraw-Hill, New York.
18. Nemhauser, G.L. and Wolsey, L.A. (1988), *Integer and Combinatorial Optimization,* Wiley, New York and Toronto.
19. Rubinov, A., Glover, B.M. and Yang, X.Q. Decreasing functions with applications to penalization, *SIAM Journal on Optimization* (to appear).
20. Rinnoy Kan, A.H.G. and Timmer, G.T. (1987), Stochastic global optimization method, *Mathematical Programming* 39: 57–78.
21. Yang, X.Q. (1998), Second order global optimality conditions for convex composite optimization, *Mathematical Programming* 81: 327–347.
22. Yang, X.Q. and Goh, C.J. A nonlinear Lagrangian function for discrete optimization problems, in Migdalas, A., Pardalos, A., Varbrand, P. and Holmqvist, K. (eds), *From Local to Global Optimization*, Kluwer Academic Publishers (to appear).