

A Continuous Approach to Binary Quadratic Problems

Zhi Liu, Zhensheng Yu, Yunlong Wang

College of Science, University of Shanghai for Science and Technology, Shanghai, China

Email: azhiliuazhi@126.com

How to cite this paper: Liu, Z., Yu, Z.S. and Wang, Y.L. (2018) A Continuous Approach to Binary Quadratic Problems. *Journal of Applied Mathematics and Physics*, 6, 1720-1732.

<https://doi.org/10.4236/jamp.2018.68147>

Received: August 3, 2018

Accepted: August 26, 2018

Published: August 29, 2018

Copyright © 2018 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This paper presents a continuous method for solving binary quadratic programming problems. First, the original problem is converted into an equivalent continuous optimization problem by using NCP (Nonlinear Complementarity Problem) function, which can be further carry on the smoothing processing by aggregate function. Therefore, the original combinatorial optimization problem could be transformed into a general differential nonlinear programming problem, which can be solved by mature optimization technique. Through some numerical experiments, the applicability, robustness, and solution quality of the approach are proved, which could be applied to large scale problems.

Keywords

Binary Quadratic Program, Continuous Approach, NCP Function, Nonlinear Programming, Aggregate Function, Multiplier Penalty Function

1. Introduction

Binary quadratic programming (BQP) problem is a kind of typical combinatorial optimization problem, and has a variety of applications in computer aided design, traffic management, cellular mobile communication frequency allocation, operations research and engineering. And many combinatorial optimization problems with constraints can be transformed into BQP problems form by certain transformation, so these problems can be on behalf of kinds of important problems in combinatorial optimization. Hammer [1] pointed out that any integer programming problems where the objective function is quadratic or linear and the constraint condition is linear can be described as BQP problems. Glover *et al.* [2] have successfully used this transformation method to transform the quadratic knapsack problem into BQP problem for solving. Considering the fea-

sibility of this transformation, BQP problem has more extensive practical application. Due to the wide application background of this problem and the difficulties caused by NP (Non-Deterministic Polynomial) hard properties, it is of great academic value to study effective algorithms for solving such problems.

Binary quadratic programming can be uniformly written in the following general

$$\begin{aligned} \min \quad & f(x) = x^T Q x \\ \text{s.t.} \quad & x_i \in \{l_i, u_i\}, i = 1, 2, \dots, n \end{aligned} \quad (1)$$

where Q is a real symmetric $n \times n$ matrix, and the objective function can be followed by another linear term, but it can be omitted because there is no substantial influence on the method introduced later. Of course, in the following discussion, we will mainly focus on the so-called 0 - 1 quadratic programming:

$$\begin{aligned} \min \quad & f(x) = x^T Q x \\ \text{s.t.} \quad & x_i \in \{0, 1\}, i = 1, 2, \dots, n \end{aligned} \quad (2)$$

However, the continuous method proposed later in this paper is also applicable to the general binary quadratic programming (1), which can be easily accomplished by simple variable transformation. It is well known that 0 - 1 programming problem will increase exponentially with the increase of the scale of the problem, and within limited computer resources and time the traditional solution method will be difficult to realize. Because of the wide application prospect and difficult characteristic of 0 - 1 programming, how to solve this kind of combinatorial optimization problem effectively has been the focus of many scholars. The classical methods for solving 0 - 1 programming problems are mainly divided into two types: accurate algorithms and random search algorithms. The accurate algorithms include implicit enumeration method, branch and bound algorithm, cutting plane algorithm, dynamic programming method, etc. Random search algorithms include genetic algorithm, simulated algorithm, artificial neural network algorithm, etc. [3] [4] [5] [6] [7]. But continuous method has become a new tendency of the 0 - 1 programming problem in recent years, whose advantage is to avoid inherent characteristics of the combinatorial optimization problem, and no longer limited to the size of the problem. This method is to transform the combinatorial optimization problem into equivalent continuous optimization problem for solving, thus effectively avoid the so-called combination explosion problem, which can solve some large combinatorial optimization problems efficiently. While this approach is still in the exploratory stage with few research work in this area, but we think it is a noteworthy new research direction. Interested readers can see the literature [8]-[13].

In this article, we will first turn the binary constraints of the 0 - 1 programming problem into equivalent complementarity problem, that is to say, a mathematical programming problem with equilibrium constraints (MPEC); then NCP function is used to transform the complementary condition into

equivalent equation. Finally, the non-smooth equation is smoothed by using the aggregate function and we finally get a nonlinear smooth optimization problem.

2. Continuous Formulation of BQP

The simplest and most intuitive way to solve 0 - 1 planning problem is to adopt the round integration technology, that is, to consider 0 - 1 variable as a continuous variable, and to replace the variable constraint $x_i \in \{0, 1\}, i = 1, 2, \dots, n$ in the original problem with the following interval constraint:

$$0 \leq x_i \leq 1 \quad (3)$$

Then a relaxation optimization problem is obtained, and the components of the relaxation solution are rounded to the nearest discrete value. Although this method is easy to implement, but the theory of this technology is not strict. It cannot guarantee that the optimal solution obtained is the global optimal solution, or even a feasible solution.

The other method is the penalty function method, which is easy to see that $x_i \in \{0, 1\}, i = 1, 2, \dots, n$ is always equivalent to a complementarity condition:

$$x_i(1 - x_i) = 0, i = 1, 2, \dots, n \quad (4)$$

Therefore, the punishment function can be constructed as follows:

$$P(\alpha, x) = \alpha \sum_{i=1}^n x_i(1 - x_i) \quad (5)$$

where $\alpha > 0$ is a large enough penalty parameter. Adding this penalty function to the object function can obtain an equivalent continuous optimization problem:

$$\min f(x) + \alpha x^T(1 - x) \quad (6)$$

Due to the strong concavity of $\alpha x^T(1 - x)$, the object function in (6) is concave. The equivalence of (2) and (6) is based on the fact that the concavity function obtains the minimum value at a vertex, as well as $x^T(1 - x) = 0$ has included $x = 0$ and $x = 1$ the sum is included. Although the vertex of the feasible domain is not necessarily the vertex of the unit hypercube, if α is sufficiently large, the global minimum can only be obtained while $x^T(1 - x) = 0$.

However, the specific value of α cannot be determined. If the value is too small, the penalty term will not play its role. If the value is too large, due to the strong concavity of the function, this penalty function method cannot effectively obtain the optimal solution of the original problem, so it is not an effective method. It is pointed out in literature [14], but only a new viewpoint is provided, and no more effective method is given to solve the original problem. Therefore, it is not ideal to solve the original problem directly with (6). Great-Ming Ng later constructed a logarithmic barrier term:

$$\Psi(x) = -\sum_{i=1}^n [\ln x_i + \ln(1 - x_i)] \quad (7)$$

This barrier term can be used to replace the boxed constraint conditions $0 \leq x_i \leq 1, i=1, 2, \dots, n$ as well as to avoid falling into the local minimum during the iteration, so as to find the global optimal solution for the original problem. At this point, the following unconstrained optimization problems can be obtained:

$$\min f(x) + \alpha x^T(1-x) - \mu \sum_{i=1}^n [\ln x_i + \ln(1-x_i)] \quad (8)$$

where $\alpha > 0$ is a penalty parameter and $\mu > 0$ is a barrier parameter. In a sense, (8) is an improvement on (6), but in essence they both replace $x_i \in \{0, 1\}, i=1, 2, \dots, n$ constraints with penalty terms or logarithmic barrier terms.

For the BQP problem, another continuous method is proposed in this paper. Consider the following two sets of constraints:

$$\begin{aligned} 0 \leq x_i \leq 1, i=1, 2, \dots, n, \\ x_i(1-x_i) = 0, i=1, 2, \dots, n. \end{aligned} \quad (9)$$

Obviously, these two sets of constraint conditions and binary constraint conditions are equivalent. In fact, these two constraints skillfully reform complementary constraints:

$$x_i \geq 0, 1-x_i \geq 0, x_i(1-x_i) = 0, i=1, 2, \dots, n \quad (10)$$

The advantage of constructing constraint conditions in this way is that: the two constraint conditions of (9) work at the same time, and to each variable there is a unique complementary constraint condition

$$x_i \geq 0, 1-x_i \geq 0, x_i(1-x_i) = 0$$

And more importantly, complementary constraints (10) and the general constraint conditions are different: for general complementary constraints, each component is a function about $x = \{x_1, x_2, \dots, x_n\}$, and for (10), the value of every $F_i(x) = x_i(1-x_i)$ only has relationship with the values of x_i , nothing to do with the other variables. Namely, each of the complementary constraint of (10) is independent of each other.

Through the above, the problem (2) can be equivalently transformed into the following mathematical programming problems with complementary constraints:

$$\begin{aligned} \min f(x) = x^T Q x \\ x_i \geq 0, 1-x_i \geq 0, x_i(1-x_i) = 0, i=1, 2, \dots, n \end{aligned} \quad (11)$$

For the above complementary constraints, we can use NCP function [15] to replace them further. To better describe NCP functions, we first introduce the concept of them:

Definition (NCP function): If a function is set $\phi: R^2 \rightarrow R$, and $\phi(a, b) = 0 \Leftrightarrow a \geq 0, b \geq 0, ab = 0$, then the function of ϕ is called an NCP function.

Therefore, we can replace the complementary constraint in (11) with a series

of equation equations. Finally, we get the continuous optimization model as follows:

$$\begin{aligned} \min \quad & f(x) = x^T Qx \\ \text{s.t.} \quad & \phi(x_i, 1 - x_i) = 0, i = 1, 2, \dots, n \end{aligned} \tag{12}$$

Thus, the global optimal solution of (12) is the exact solution of problem (2).

3. Smoothing Method for BQP

According to the definition of NCP function given above, functions satisfying conditions can take various forms. Several commonly used NCP functions can be given:

$$\begin{aligned} \phi_M &= \min(a, b) \\ \phi_{FB} &= \sqrt{a^2 + b^2} - (a + b) \\ \phi(a, b) &= -ab + \frac{1}{2} \min^2\{0, a + b\} \\ \phi(a, b) &= (a - b)^2 - a|a| - b|b| \end{aligned} \tag{13}$$

In the following section, we mainly choose the minimum value function to study.

It is easy to know $\min\{a, b\} = -\max\{-a, -b\}$, and for the non-differentiable maximum function:

$$f_{\max}(x) = \max\{f_i(x)\}, i = 1, 2, \dots, n \tag{14}$$

we often use aggregate function to carry out smooth approximation to it [16] [17]. The so-called aggregate function is a function in the following form:

$$F_\mu(x) = \mu^{-1} \left\{ \ln \sum_{i=1}^m \exp[\mu f_i(x)] \right\} \tag{15}$$

where the smooth parameter μ is large enough, and $f_i(x), i = 1, 2, \dots, n$ are continuous differentiable (i.e., smooth) real value function. The aggregate function has very good properties [18], which can be found from the following properties:

Lemma 3.1 *The function $F_\mu(x)$ which is defined by (15) satisfies the following conditions:*

- 1) $f_{\max}(x) < F_\mu(x) \leq f_{\max}(x) + \mu^{-1} \ln m$
- 2) $\lim_{\mu \rightarrow \infty} F_\mu(x) = f_{\max}(x)$

Proof 1) $F_\mu(x)$ can be equivalent to deformation $F_\mu(x) = f_{\max}(x) + \mu^{-1} \ln \sum_{i=1}^m \exp\{\mu[f_i(x) - f_{\max}(x)]\}$, because $f_i(x) \leq f_{\max}(x)$, so $0 < \exp\{\mu[f_i(x) - f_{\max}(x)]\} \leq 1$, and there is at least one indicator that makes $f_i(x) = f_{\max}(x)$, so exist $\exp\{\mu[f_i(x) - f_{\max}(x)]\} = 1$. So we have

$$0 < f_i(x) - f_{\max}(x) \leq \mu^{-1} \ln m \tag{16}$$

- 2) When $\mu \rightarrow \infty$, the above conclusion can be proved by formula (16).

Based on the above introduction, we can handle the minimum function as follows [19]:

$$\begin{aligned}\phi_M(a, b) &= \min(a, b) = -\max(-a, -b) \approx -F_\mu(-a, -b) \\ &= -\mu^{-1} \ln[\exp(-\mu a) + \exp(-\mu b)]\end{aligned}\quad (17)$$

where the smooth parameter μ is large enough, as shown by lemma 3.1, when $\mu \rightarrow \infty$, $-F_\mu(-a, -b)$ and $\phi_M(a, b)$ are equivalent.

For ease of expression, let's call:

$$\begin{aligned}\phi_M(x_i, 1-x_i) &= \min\{x_i, 1-x_i\}, i=1, 2, \dots, n \\ \phi_\mu(x_i, 1-x_i) &= \mu^{-1} \ln[\exp(-\mu x_i) + \exp(-\mu(1-x_i))], i=1, 2, \dots, n\end{aligned}\quad (18)$$

At this point, the problem (12) can be further transformed into an equivalent continuous and smooth nonlinear constraint optimization problem:

$$\begin{aligned}\min \quad & f(x) = x^T Qx \\ \text{s.t.} \quad & \phi_\mu(x_i, 1-x_i) = 0, i=1, 2, \dots, n\end{aligned}\quad (19)$$

By this continuous method, the global optimal solution of the original problem (2) can be solved by solving the corresponding nonlinear constrained optimization problem (19). Many mature optimization algorithms have been developed for solving nonlinear equality constraint optimization problems. This paper mainly uses the multiplier penalty function method to solve the above problems (19). Multiplier method is an optimization algorithm independently proposed by Powell and Hestenes in 1969 to solve equality constraint optimization problems. Later, it was extended to solve inequality constraint optimization problems by Rockfellar in 1973. The basic idea is to start from the Lagrangian function of the original problem and add an appropriate penalty function, so as to transform the original problem into a series of unconstrained optimization subproblems.

For convenience's sake, let's first:

$$\begin{aligned}\Phi(x) &= (\phi_M(x_1, 1-x_1), \phi_M(x_2, 1-x_2), \dots, \phi_M(x_n, 1-x_n))^T \\ \Phi(x, \mu) &= (\phi_\mu(x_1, 1-x_1), \phi_\mu(x_2, 1-x_2), \dots, \phi_\mu(x_n, 1-x_n))^T\end{aligned}\quad (20)$$

The Lagrange function of problem (19) is [20]:

$$L(x, \lambda) = f(x) + \lambda^T \Phi(x, \mu) \quad (21)$$

where $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)^T$ is the Lagrangian multiplier vector, and (x^*, λ^*) is set as KT pairs of problem (19), then it can be known from the optimality condition:

$$\nabla_x L(x^*, \lambda^*) = 0, \quad \nabla_\lambda L(x^*, \lambda^*) = \Phi(x^*, \mu) = 0 \quad (22)$$

In addition, it is not difficult to find any x in the feasible domain is satisfied:

$$L(x^*, \lambda^*) = f(x^*) \leq f(x) = f(x) + (\lambda^*)^T \Phi(x, \mu) = L(x, \lambda^*) \quad (23)$$

The above equation shows that if the multiplier vector λ^* is known, the

problem (19) can be equivalently transformed into:

$$\begin{aligned} \min \quad & L(x, \lambda^*) \\ \text{s.t.} \quad & \phi_\mu(x_i, 1 - x_i) = 0, i = 1, 2, \dots, n \end{aligned} \tag{24}$$

Then the external penalty function method is considered to solve the problem (24). The augmented objective function is

$$L_\mu(x, \lambda, \alpha) = L(x, \lambda) + \frac{\alpha}{2} \|\Phi(x, \mu)\|^2 = f(x) + \lambda^T \Phi(x, \mu) + \frac{\alpha}{2} \|\Phi(x, \mu)\|^2 \tag{25}$$

In this way, we can fix $\lambda = \bar{\lambda}$ and find a minimum of $L_\mu(x, \bar{\lambda}, \alpha)$, then change the value of λ appropriately to find a new value \bar{x} , until we get the x^* and λ^* that we want. Specifically, when solving the minimum $x^{(k)}$ of $\min L_\mu(x, \lambda^{(k)}, \alpha)$ in the k-th iteration of the unconstrained subproblem, the necessary conditions for taking the extreme values are known:

$$\nabla_x L_\mu(x^{(k)}, \lambda^{(k)}, \alpha) = \nabla f(x^{(k)}) + \nabla \Phi(x^{(k)}, \mu) [\lambda^{(k)} + \alpha \Phi(x^{(k)}, \mu)] = 0 \tag{26}$$

And the KT-point of (x^*, λ^*) the original problem satisfies:

$$\nabla f(x^*) + \nabla \Phi(x^*, \mu) \lambda^* = 0, \quad \Phi(x^*, \mu) = 0 \tag{27}$$

In order for $\{x^{(k)}\} \rightarrow x^*$ and $\{\lambda^{(k)}\} \rightarrow \lambda^*$, after comparing the above two expressions, the updating formula of the multiplier sequence $\{\lambda^{(k)}\}$ is:

$$\lambda_i^{(k+1)} = \lambda_i^{(k)} + \alpha^{(k)} \phi_{\mu^{(k)}}(x_i^{(k)}, 1 - x_i^{(k)}), i = 1, 2, \dots, n \tag{28}$$

It can be seen from Equation (28) that the sufficient and necessary condition for $\{\lambda^{(k)}\}$ convergence is $\{\phi_\mu(x^{(k)}, 1 - x^{(k)})\} \rightarrow 0$. And then we proof that $\phi_\mu(x^{(k)}, 1 - x^{(k)}) = 0$ is also a necessary and sufficient condition for judging the KT-point (x^*, λ^*) .

Theorem 3.1 Let $x^{(k)}$ be the minimum point of the unconstrained optimization problem:

$$\min L_\mu(x, \lambda^{(k)}, \alpha) = L(x, \lambda^{(k)}) + \frac{\alpha}{2} \|\Phi(x, \mu)\|^2 \tag{29}$$

then $(x^{(k)}, \lambda^{(k)})$ being a KT-point to the (19) has a necessary and sufficient condition of $\phi_\mu(x^{(k)}, 1 - x^{(k)}) = 0$.

Proof The necessity is obvious. The following proof is sufficient, since $x^{(k)}$ is a minimum of (29) and $\phi_\mu(x^{(k)}, 1 - x^{(k)}) = 0$, therefore, for any feasible point:

$$f(x) = L_\mu(x, \lambda^{(k)}, \alpha) \geq L_\mu(x^{(k)}, \lambda^{(k)}, \alpha) = f(x^{(k)}) \tag{30}$$

that is, $x^{(k)}$ is also a minimum of (19). On the other hand, it is noted that $x^{(k)}$ is also the stable point of (29), therefore

$$\nabla_x L_\mu(x^{(k)}, \lambda^{(k)}, \alpha) = \nabla f(x^{(k)}) + \nabla \Phi(x^{(k)}, \mu) \lambda^{(k)} = 0 \tag{31}$$

The above formula indicates that $\lambda^{(k)}$ is the Lagrangian multiplier vector with respect to $x^{(k)}$, that is to say $(x^{(k)}, \lambda^{(k)})$ is also the KT-point of (19).

Based on the above discussion, before giving the detailed steps to solve the

multiplier method of equation constraint problem (19), the following lemmas are given by studying some properties of $\phi_\mu(x_i, 1-x_i)$.

Lemma 3.2 For each $i = 1, 2, \dots, n$, the following formula holds:

$$\phi_M(x_i, 1-x_i) - \mu^{-1} \ln 2 \leq \phi_\mu(x_i, 1-x_i) \leq \phi_M(x_i, 1-x_i) \tag{32}$$

Proof According to the properties of aggregate function:

$$f_{\max}(-x_i, -1+x_i) \leq F_\mu(x_i, 1-x_i) \leq f_{\max}(-x_i, -1+x_i) + \mu^{-1} \ln 2 \tag{33}$$

thus $-f_{\max}(-x_i, -1+x_i) - \mu^{-1} \ln 2 \leq \phi_\mu(x_i, 1-x_i) \leq -f_{\max}(-x_i, -1+x_i)$. That is to say, $\min\{x_i, 1-x_i\} - \mu^{-1} \ln 2 \leq \phi_\mu(x_i, 1-x_i) \leq \min\{x_i, 1-x_i\}$. Therefore, $\phi_M(x_i, 1-x_i) - \mu^{-1} \ln 2 \leq \phi_\mu(x_i, 1-x_i) \leq \phi_M(x_i, 1-x_i)$.

Lemma 3.3 For any $\mu > 0$, there is

$$\begin{aligned} |\phi_M(x_i, 1-x_i) - \phi_\mu(x_i, 1-x_i)| &\leq \mu^{-1} \ln 2 \\ \|\Phi(x) - \Phi(x, \mu)\| &\leq \mu^{-1} \sqrt{n} \ln 2 \end{aligned} \tag{34}$$

Proof Easy to obtain by lemma 3.2 that:

$0 \leq \phi_M(x_i, 1-x_i) - \phi_\mu(x_i, 1-x_i) \leq \mu^{-1} \ln 2$, so

$$|\phi_M(x_i, 1-x_i) - \phi_\mu(x_i, 1-x_i)| \leq \mu^{-1} \ln 2, i = 1, 2, \dots, n. \tag{35}$$

From the above equation we know:

$$\begin{aligned} 0 \leq \|\Phi(x) - \Phi(x, \mu)\| &= \left\{ \sum_{i=1}^n [\phi_M(x_i, 1-x_i) - \phi_\mu(x_i, 1-x_i)]^2 \right\}^{\frac{1}{2}} \\ &\leq \left[(\mu^{-1} \ln 2)^2 n \right]^{\frac{1}{2}} = \mu^{-1} \sqrt{n} \ln 2 \end{aligned} \tag{36}$$

Lemma 3.4 For each $i = 1, 2, \dots, n$, $\phi_\mu(x_i, 1-x_i)$ is strictly convex on the interval $(-\infty, +\infty)$; When $\mu \geq 10^2$, $\phi_\mu^2(x_i, 1-x_i)$ is convex on the interval $(-\infty, 0.4737) \cup (0.5263, +\infty)$.

Proof According to the definition of $\phi_\mu(x_i, 1-x_i)$, it can be seen that:

$$\begin{aligned} \phi_\mu(x_i, 1-x_i) &= \mu^{-1} \ln \left\{ \exp(-\mu x_i) + [-\mu(1-x_i)] \right\} \\ \phi_\mu^2(x_i, 1-x_i) &= \left\langle \mu^{-1} \ln \left\{ \exp(-\mu x_i) + [-\mu(1-x_i)] \right\} \right\rangle^2 \end{aligned} \tag{37}$$

let $r_1 = e^{-\mu x_i} / (e^{-\mu x_i} + e^{-\mu(1-x_i)})$, $r_2 = e^{-\mu(1-x_i)} / (e^{-\mu x_i} + e^{-\mu(1-x_i)})$. We could get,

$$\begin{aligned} \phi_\mu''(x_i, 1-x_i) &= 4\mu r_1 r_2 > 0, i = 1, 2, \dots, n; \\ (\phi_\mu^2(x_i, 1-x_i))'' &= 8\mu \left[(r_1 - r_2)^2 / (4\mu r_1 r_2) + \phi_\mu(x_i, 1-x_i) \right]. \end{aligned} \tag{38}$$

So $\phi_\mu(x_i, 1-x_i)$ is strictly convex on the interval $(-\infty, +\infty)$; for $\phi_\mu^2(x_i, 1-x_i)$, when $\mu \geq 10^2$ all the $x \notin (0.4737, 0.5263)$ satisfy $(\phi_\mu^2(x_i, 1-x_i))'' > 0$, so when $\mu \geq 10^2$, $\phi_\mu^2(x_i, 1-x_i)$ is convex on the interval $(-\infty, 0.4737) \cup (0.5263, +\infty)$.

Lemma 3.5 For any $\mu > 0$ and $\alpha > 0$ that are large enough, as long as the region where x satisfies the following equation:

$(r_2 - r_1)^2 / (4\mu r_1 r_2) + \phi_\mu(x_i, 1-x_i) > 0$, the augmented Lagrange function $L_\mu(x, \lambda, \alpha)$ defined by Equation (25) is convex in this region.

Proof Let $l(x, \lambda) = x^T Qx + \lambda^T \Phi(x, \mu)$ then the hesse matrix of $L_\mu(x, \lambda, \alpha)$ is

$$\nabla_{xx}^2 L_\mu(x, \lambda, \alpha) = \nabla_{xx}^2 l(x, \lambda) + \alpha B(x, p) \tag{39}$$

where

$$B(x, p) = \text{Diag} \left[\left(\phi'_\mu(x_i, 1-x_i) \right)^2 + \phi_\mu(x_i, 1-x_i) \phi''_\mu(x_i, 1-x_i) \right], i = 1, 2, \dots, n \tag{40}$$

and

$$\begin{aligned} & \left(\phi'_\mu(x_i, 1-x_i) \right)^2 + \phi_\mu(x_i, 1-x_i) \phi''_\mu(x_i, 1-x_i) \\ & = 8\mu \left[(r_2 - r_1)^2 / (4\mu r_1 r_2) + \phi_\mu(x_i, 1-x_i) \right] \end{aligned} \tag{41}$$

So we just have to satisfy $(r_2 - r_1)^2 / (4\mu r_1 r_2) + \phi_\mu(x_i, 1-x_i) > 0$, the matrix $B(x, p)$ is positive, therefore for sufficiently large $\alpha > 0$, $L_\mu(x, \lambda, \alpha)$ is convex in the region that satisfies the above conditions. When $\mu > 10^3$ the interval is $\phi_\mu^2(x_i, 1-x_i)$ is convex on the interval $(-\infty, 0.4962) \cup (0.5038, +\infty)$; when $\mu > 10^4$ the interval is $(-\infty, 0.4995) \cup (0.5005, +\infty)$.

4. Algorithm

According to lemma 3.5, we can use the augmented Lagrange penalty function to solve the problem (19). For a strictly monotonic increasing sequence $\{\mu^k\}$ and $\{\alpha^k\}$, the solution of unconstrained optimization $\min_{x \in \mathbb{R}^n} L_\mu(x, \lambda^{(k)}, \alpha^{(k)})$ is $\{x^k\}$, the corresponding Lagrange multiplier is $\{\lambda^k\}$, and is modified according to (28) in the iteration.

Based on the previous analysis, the basic algorithm for solving the problem (19) is as follows [19]:

Algorithm 1

Step 1 Given parameters $\mu^{(0)} > 0$, $\alpha^{(0)} > 0$, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$ and $\sigma_1 > 1$, $\sigma_2 > 1$. Select a starting point $x^{(s,0)}, \lambda^{(0)}$ and set $t^{(0)} = \|\Phi(x^{(s,0)}, \mu^{(0)})\|$, $k = 0$;

Step 2 Using BFGS algorithm solve the unconstrained minimization problem $\min_{x \in \mathbb{R}^n} L_\mu(x, \lambda^{(k)}, \alpha^{(k)})$ with the starting point $x^{(s,k)}$, and denote by $x^{(k)}$ its optimal solution;

Step 3 If $\|\Phi(x^{(k)}, \mu^{(k)})\| \leq \varepsilon_1, \|f(x^{(k)}) - f(x^{(s,k)})\| \leq \varepsilon_2$, set $x^* = x^{(k)}$ go to **Step 6**; else go to **Step 4**;

Step 4 If $\|\Phi(x^{(k)}, \mu^{(k)})\| \leq 0, 1t^{(k)}$ go to **Step 5**; else update the parameter $\alpha^{(k+1)} = \sigma_1 \alpha^{(k)}$, $\mu^{(k+1)} = \sigma_2 \mu^{(k)}$ set $k = k + 1$, and go to **Step 1**;

Step 5 Set $t^{(k)} = \|\Phi(x^{(k)}, \mu^{(k)})\|$, $x^{(s,k+1)} = x^{(k)}$, modify $\lambda^{(k+1)}$ according to Equation (28), set $k = k + 1$, and go to **Step 1**;

Step 6 Finish.

The optimal solution sequence $x^{(k)}$ required by the above algorithm makes $\|\Phi(x^{(k)}, \mu^{(k)})\|$ converge to 0 at a speed of at least 0.1. If this requirement is not met in an iteration, the penalty factor and smooth parameter should be increased automatically.

Then we introduce the BFGS algorithm mentioned in step 2. It is the most popular and effective quasi-newton methods at present. It was proposed by Broyden, Fletcher, Goldfarb, and Shanno independently in 1970, so it is called BFGS algorithm. We know that the basic idea of quasi-newton method is to replace the Hesse matrix $G^{(l)} = \nabla^2 f(x^{(l)})$ with one of its approximate matrixes $B^{(l)}$ in the steps of basic Newton method, and the following relational expression is required:

$$B^{(l+1)}s^{(l)} = y^{(l)} \quad (42)$$

where displacement $s^{(l)} = x^{(s,l)} - x^{(s,l-1)}$, gradient difference $y^{(l)} = g^{(l)} - g^{(l-1)}$, (42) is usually referred to as quasi-newton equation or quasi-newton condition.

Algorithm 2 (BFGS)

Step 1 Given parameters $\delta \in (0,1)$, $\sigma \in (0,0.5)$ select a starting point $x^{(s,l)}$ and positive matrices $B^{(0)}$ (usually set as $G(x_0)$ or the unit matrix I_n), and set $l = 0$, $f(x) = L_\mu(x, \lambda^{(k)}, \alpha^{(k)})$;

Step 2 Calculate $g^{(l)} = \nabla f(x^{(s,l)})$, if $\|g^{(l)}\| \leq \varepsilon$ go to **Step 6**, output $x^{(s,l)}$ as approximate minimum point; else go to **Step 3**;

Step 3 Solve the linear equations $B^{(l)}d = -g^{(l)}$ and get the solution $d^{(l)}$;

Step 4 Suppose $m^{(l)}$ is the minimum non-negative solution that satisfies the following inequality:

$$f(x^{(s,l)}) + \delta^m d^{(l)} \leq f(x^{(s,l)}) + \sigma \delta^m (g^{(l)})^T d^{(l)}$$

set $\delta^{(l)} = \delta^{m^{(l)}}$, $x^{(s,l+1)} = x^{(s,l)} + \alpha^{(l)} d^{(l)}$;

Step 5 Set $s^{(l)} = x^{(s,l)} - x^{(s,l-1)}$ $y^{(l)} = g^{(l)} - g^{(l-1)}$, by the correction formula:

$$B^{(l+1)} = \begin{cases} B^{(l)} & (y^{(l)})^T s^{(l)} \leq 0 \\ B^{(l)} - \frac{B^{(l)}s^{(l)}(s^{(l)})^T B^{(l)}}{(s^{(l)})^T B^{(l)}s^{(l)}} + \frac{y^{(l)}(y^{(l)})^T}{(y^{(l)})^T s^{(l)}} & (y^{(l)})^T s^{(l)} > 0 \end{cases} \quad (43)$$

identify the $B^{(l+1)}$, set $l = l + 1$ and go to **Step 1**;

Step 6 Finish.

5. Number Experiments

About BQP questions, the current relatively popular in the question bank is J. E. Beasley's OR-library

(<http://people.brunel.ac.uk/~%20mastjib/jeb/orlib/bqpinf.html>), in this paper, we has carried on the numerical experiment on parts of question bank BQP questions, and the calculation results of our continuous algorithm are compared with those of the best solution of known, as shown in **Table 1**.

The first column Mu in the **Table 1** is the question number; The second column m denotes the number of variables; The third is the numerical result of this paper; The fourth is the best known solution; The fifth column represents our numerical result as a percentage of the best solution.

Table 1. Comparison of numerical results.

Mu	m	solution	best solution	percent	Mu	m	solution	best solution	percent
1a	50	3414	3414	100	4d	100	10,616	10,727	98.97
2a	60	6063	6063	100	5d	100	10,960	11,626	94.27
3a	70	5992	6037	99.25	6d	100	13,689	14,207	96.35
4a	80	8589	8598	99.9	7d	100	14,080	14,476	97.26
5a	50	5377	5737	93.72	8d	100	16,033	16,352	98.05
6a	30	3973	3980	99.82	9d	100	15,584	15,656	99.54
7a	30	4497	4541	99.03	10d	100	18,952	19,102	99.21
8a	100	11109	11109	100	1e	200	16,254	16,464	98.72
1c	40	5058	5058	100	2e	200	23,064	23,395	98.59
2c	50	6213	6213	100	3e	200	24,514	25,243	97.11
3c	60	6665	6665	100	4e	200	35,283	35,594	99.13
4c	70	7221	7398	97.61	5e	200	33,577	35,154	95.51
5c	80	7326	7362	99.51	1f	500	59,661	61,194	97.49
6c	90	5610	5824	96.33	2f	500	98,129	100,161	97.97
7c	100	7163	7225	99.14	3f	500	133,524	138,035	96.73
1d	100	6157	6333	97.22	4f	500	169,606	172,771	98.17
2d	100	6444	6579	97.95	5f	500	187,476	190,507	98.41
3d	100	9170	9261	99.02					

It's not hard to see from the **Table 1** that in the process of solving these problems, the optimal solution obtained by our algorithm is basically close to the best known solution. Through the comparison in the table above, it can be seen that the proposed continuous algorithm is basically close to the known best solution for solving the general unconstrained 0 - 1 quadratic programming problem.

6. Conclusion

In this paper, we have reformulated an unconstrained BQP problem as a MPEC problem by the equivalent complementarity conditions of a binary vector. To seek a global minimizer of the resulting continuous optimization problem, we construct a global smoothing function and develop a global continuation algorithm via a sequence of unconstrained minimization. And the numerical results indicate that the continuous approach proposed is extremely promising, especially for those large problems, in terms of the quality of the optimal values generated and the computational work involved. In addition, this new approach can be extended to general nonlinear binary optimization problems, and we can

leave it as a future research topic.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Hammer, P. and Rudeanu, S. (1968) Boolean Methods in Operations Research. Elsevier B.V., New York. <https://doi.org/10.1007/978-3-642-85823-9>
- [2] Glover, F., Kochenberger, G., Alidaee, B. and Amini, M.M. (1999) Tabu with Search Critical Event Memory: An Enhance Application for Binary Quadratic Programs. Kluwer Academic Publisher, Boston.
- [3] Goffin, J.L. (1997) Solving Nonlinear Multicommodity Flow Problems by the Analytic Center Cutting Plane Method. *Mathematical Programming*, **76**, 131-154. <https://doi.org/10.1007/BF02614381>
- [4] Christoph, H. and Rendl, F. (1998) Solving Quadratic (0,1)-Problems by Semidefinite Programs and Cutting Planes. *Mathematical Programming*, **82**, 291-315. <https://doi.org/10.1007/BF01580072>
- [5] Wolkowicz, H. and Anjos, M.F. (2002) Semidefinite Programming for Discrete Optimization and Matrix Completion Problem. *Discrete Applied Mathematics*, **123**, 513-577. [https://doi.org/10.1016/S0166-218X\(01\)00352-3](https://doi.org/10.1016/S0166-218X(01)00352-3)
- [6] Beasley, J.E. (1998) Heuristic Algorithms for the Unconstrained Binary Quadratic. Management School of Imperial College of U.K., London, 1-36.
- [7] Glover, F. (2002) One-Pass Heuristics for Large-Scale Unconstrained Binary Quadratic Problems. *European Journal of Operational Research*, **137**, 272-287. [https://doi.org/10.1016/S0377-2217\(01\)00209-0](https://doi.org/10.1016/S0377-2217(01)00209-0)
- [8] Warners, J.P. (1997) Potential Reduction Algorithms for Structured Combinatorial Optimization Problems. *Operations Research Letters*, **21**, 55-64. [https://doi.org/10.1016/S0167-6377\(97\)00031-X](https://doi.org/10.1016/S0167-6377(97)00031-X)
- [9] Audet, C. (1997) Links between Linear Bilevel and Mixed 0 - 1 Programming Problems. *Journal of Optimization Theory and Applications*, **93**, 273-300. <https://doi.org/10.1023/A:1022645805569>
- [10] Kiwiel, K.C. (2000) Bregman Proximal Relaxation of Large-Scale 0 - 1 Problems. *Computational Optimization and Applications*, **15**, 33-44. <https://doi.org/10.1023/A:1008770914218>
- [11] Pardalos, P.M. (2000) Recent Developments and Trends in Global Optimization. *Journal of Computational and Applied Mathematics*, **124**, 209-228. [https://doi.org/10.1016/S0377-0427\(00\)00425-8](https://doi.org/10.1016/S0377-0427(00)00425-8)
- [12] Pardalos, P.M. (1996) Continuous Approaches to Discrete Optimization Problems. Nonlinear Optimization and Applications. Plenum Publishing, New York, 313-328. https://doi.org/10.1007/978-1-4899-0289-4_22
- [13] Ng, K.M. (2002) A Continuation Approach for Solving Nonlinear Optimization Problems with Discrete Variables. Stanford University, Stanford.
- [14] Horst, R., Pardalos, P.M. and Thoai, N.V. (1995) Introduction to Global Optimization. Kluwer Academic Publisher, Boston.
- [15] Mangasarian, O.L. (1976) Equivalence of the Complementarity Problem to a System of Nonlinear Equations. *SIAM Journal on Applied Mathematics*, **31**, 89-92.

<https://doi.org/10.1137/0131009>

- [16] Li, X. (1991) An Aggregate Constrained Method for Nonlinear Programming. *The Operations Research Society*, **42**, 1003-1010. <https://doi.org/10.1057/jors.1991.190>
- [17] Li, X. (1992) An Entropy-Based Aggregate Method for Minimax Optimization. *Engineering Optimization*, **18**, 277-285. <https://doi.org/10.1080/03052159208941026>
- [18] Li, Y. and Li, X. (2009) Continuous Approaches to 0 - 1 Programming Problem with Applications. PhD Thesis, Dalian University of Technology, Dalian.
- [19] Tao, T. and Li, X. (2006) Continuous Approaches to Discrete Optimum Design. PhD Thesis, Dalian University of Technology, Dalian.
- [20] Ma, C. (2010) Optimization Method and Matlab Program Design. Science Press, Beijing.