

Modified Courant-Beltrami penalty function and a duality gap for invex optimization problem

Mansur Hassan^{1,2,*} and Adam Baharum¹

¹ School of Mathematical Sciences, Main Campus, USM, 11800 Geligor, Pulau Penang, Malaysia

² Department of Mathematics, Yusuf Maitama Sule University, Kano, 700241 Kabuga, Kano, Nigeria

Received: 1 November 2018 / Accepted: 13 May 2019

Abstract. In this paper, we modified a Courant-Beltrami penalty function method for constrained optimization problem to study a duality for convex nonlinear mathematical programming problems. Karush-Kuhn-Tucker (KKT) optimality conditions for the penalized problem has been used to derived KKT multiplier based on the imposed additional hypotheses on the constraint function g . A zero-duality gap between an optimization problem constituted by invex functions with respect to the same function η and their Lagrangian dual problems has also been established. The examples have been provided to illustrate and proved the result for the broader class of convex functions, termed invex functions.

Keywords: Courant-Beltrami penalty function / penalized problem / Lagrangian dual

1 Introduction

Duality principle provided an avenue to view an optimization problem from one of the two perspectives, the primal problem (original formulation) or the dual problem (an alternative formulation). The purpose of duality in mathematical programming is to devise an alternative to the optimization problem which might be more convenient computationally or analytically. One interesting property of the dual problem is that its solution need not be equal to that of the primal problem. However, it provides a lower bound to the solution of the primal minimization problem. In general, the difference between the optimal values of the two problems is term as duality gap, if the optimal coincide then no duality gap occurs.

Based on the recent development, one of the suitable approaches in solving nonlinear optimization problem is an unconstrained optimization technique, this technique by extension can also be used to solve constrained optimization problem via a penalty function method. The idea is to carry out by transforming the original nonlinear constrained optimization problem to a single or a sequence of unconstrained optimization problems [1].

Several types of researches have been conducted to demonstrate the idea of a penalty function method and Lagrangian duality in contrast to the conventional approach for solving the constrained optimization problem [2-9].

The concept of duality and penalty function method are essential topics in optimization theory, methods and its applications [10]. There are many advancements regarding those topics in mathematical programming. The work of Changyu et al. [11] proposes a general dual program via generalized nonlinear Lagrangian functions for the constrained optimization problem. However, a dual program consists of a class of general dual programs that possess some explicit structures as special cases. Gao et al. [12] justified that a dual canonical transformation can be used to solve the general non-convex quadratic minimization problem with non-convex constraints, this can only be achieved by converting a non-convex problem into a canonical dual problem with zero duality gap. Antczak [13] combined the idea of exact penalty function and duality theory to investigate and prove the zero duality gap between optimization problems constituted by invex functions introduced by Hanson [14] named by Craven [15] and their Lagrangian dual under some mild assumptions. Ernst and Volle [16] pay more attention to conic convex and established a general zero duality gap for conic convex via generalized Courant-Beltrami penalty function. Kanzi and Soleimani-damaneh [17] proposed and reviewed the Slater constraint qualification (CQ) specifically for a semi-infinite optimization problem, which involved upper semi-continuous quasi-convex objective and constraint functions, Karush-Kuhn-Tucker (KKT) necessary and sufficient optimality conditions were provided.

The distance function was used by Soleimani-Damaneh [18] to illustrates a penalization mechanism, it was specifically designed for providing necessary and sufficient

* e-mail: mansurkust@yahoo.com

conditions to the solutions of the variational inequality problem. Moreover, Soleimani-Damaneh [19] used mordukhovich's subdifferential to solve nonsmooth optimization and variational problems.

In this paper, we modified a Courant-Beltrami penalty function method, prove and obtain a few optimality and duality results. Precisely, for smooth optimization problems involving invex functions with respect to the same function η via a modified Courant-Beltrami (MCB) penalty function method. Lagrange multiplier vectors based on MCB penalty has been derived. Moreover, a zero-duality gap under some mild assumptions was also established and validated with some examples.

The entire presentation is organized as follows: In Section 2, some preliminary definitions and results that will be useful in the subsequent sections were presented. In Section 3, we present the concept of a penalty function approach and the modification made to Courant-Beltrami penalty function. In Section 4, the derivation of KKT multiplier in view of MCB penalty was presented. In Section 5, the Lagrangian duality and some examples were discussed and finally, we conclude the paper.

2 Preliminary result

We consider the following optimization problem:

$$\begin{aligned} & \text{minimize} && f(x) \\ & \text{Subject to} && g_i(x) \leq 0, i \in I = \{1, 2, \dots, m\} \\ & && x \in R^n \end{aligned} \quad (1)$$

where $f: X \rightarrow R$ and $g_i: X \rightarrow R$, $i \in I$, are continuously differentiable functions on a nonempty set $X \subset R^n$.

Some notations that are used throughout this presentation will be introduced. Let $F = \{x \in R^n : g_i(x) \leq 0, \forall i = 1, 2, \dots, m\}$, be the set of all feasible solutions of the problem (1).

If the set of indices I partition into I_1 and I_2 with $I_1 \cap I_2 = \emptyset$, $I_1 \cup I_2 = I$, the set of active constraints at a point $x \in R^n$ can be denoted as $I_1(\bar{x}) = \{i \in I : g_i(\bar{x}) = 0\}$.

Definition 1. [13] If the set of all feasible solutions in the problem (1) is not empty, we say that problem (1) is consistent.

Definition 2. [13] If there exists a strictly feasible solution \bar{x} , that is, $g_i(\bar{x}) < 0$, $i \in I$, then problem (1) is super-consistent, and the feasible point \bar{x} is called a Slater point for the problem (1).

Under slightly reasonable convexity assumptions of $f(x)$, and $g_i(x)$, the KKT conditions are also necessary and sufficient for optimality of such optimization problem (1).

Theorem 1. Let \bar{x} be an optimal solution to the problem (1) and the suitable constraint qualification [10] be satisfied at \bar{x} . Then there exist the Lagrange multipliers $\bar{\xi}_i$, $i = 1, \dots, m$ such that

$$\nabla f(\bar{x}) + \sum_{i=1}^m \bar{\xi}_i \nabla g_i(\bar{x}) = 0 \quad (2)$$

$$\bar{\xi}_i g_i(\bar{x}) = 0, i = 1, \dots, m \quad (3)$$

$$\bar{\xi}_i \geq 0, i = 1, \dots, m. \quad (4)$$

Let assume that \bar{x} is an optimal solution to mathematical programming problem (1) and a suitable constraint qualification be satisfied. Furthermore, the necessary optimality of KKT conditions holds at \bar{x} with the optimal vector of KKT multiplier $\xi \in R^m$. Then the Lagrangian can be defined as

$$L(x, \xi) = f(x) + \xi g(x). \quad (5)$$

Invex function is the wider class of convex introduced by Hanson [14] and named by Craven [15]. This new class of function is termed generalized convex functions.

Definition 3. [14] Let $f: X \rightarrow R$ be a differentiable function on $X \subset R^n$ and $u \in R^n$. If, there exists a vector-valued function $\eta: R^n \times R^n \rightarrow R^n$ such that, $\forall x \in X$, the following inequality

$$f(x) - f(u) \geq \nabla f(u) \eta(x, u) \quad (>). \quad (6)$$

Holds, then the function f is said to be an invex (strictly invex) function with respect to η at u on X .

If (6) holds at each point $u \in R^n$, then f is said to be an invex (strictly invex) function with respect to η on R^n .

For a problem (1) to be an invex optimization problem, all functions constituting the problem (1) must be an invex function on R^n with respect to the same function η . If the optimization problem (1) is super-consistent with an invex constraint g_i on R^n (w.r.t. the same function η) then the problem (1) satisfies the Slater constraint qualification (CQ).

3 Penalty function

Penalty method is one of the suitable approaches for solving nonlinear constrained optimization problems. The idea is achievable if constrained optimization is replaced by a single or a sequence of unconstrained problems. Ideally, the solutions of the transformed problem will always converge to the solution of the original problem. The primary role of the penalty term is to penalize the problem whenever a constraint is violated and to search for the optimum among the feasible points. There are several penalty functions in the literature, and the most popular among them was the absolute value penalty function introduced by Zangwill [1]. For a given constraint $g_i(x) \leq 0$, the absolute value penalty function is

$$p(x) = \sum_{i=1}^m g_i^+(x). \quad (7)$$

Note that the function $g_i^+(x)$ is defined by $g_i^+(x) = \max\{0, g_i(x)\}$, penalty function (7) was later modified to quadratic form, popularly known as

Courant-Beltrami penalty function [8,9]

$$p(x) = \sum_{i=1}^m \{g_i^+(x)\}^2. \tag{8}$$

Definition 4. [1] A continuous function $p: R^n \rightarrow R$, is said to be a generalized penalty function (penalty function) for the constrained optimization problem (1) if it satisfied the following conditions:

- $p(x) = 0$ for all feasible points x to the problem (1)
- $p(x) > 0$ for infeasible points to the problem (1)

Note, that the conditions in Definition 4 are direct interpretations of whether the constraints function g_i is violated or not.

We modified a quadratic form of penalty function (8), popularly known as Courant-Beltrami penalty function, to a logarithmic form based on the new logarithmic penalty function that was restricted to equality constraints introduced by Hassan and Baharum [20]. The modified Courant-Beltrami (MCB) penalty function performs the same role as the existing penalty functions. The MCB penalty function can be express in the following form:

$$p(x) = \sum_{i=1}^m \ln\left(\left(g_i^+(x)\right)^2 + 1^i\right), \quad i \in I = \{1, 2, \dots, m\}. \tag{9}$$

Let a non-negative constant c be a penalty parameter, $c_k \geq 0, k = 1, 2, \dots$ may be a sequence of penalty parameters with $c_{k+1} \geq c_k \forall k$ and $\lim_{k \rightarrow \infty} c_k = \infty$.

Accordingly, we converted the considered optimization problem (1) into the single of unconstrained problems based on the penalty function (9) as follows

$$\begin{aligned} \min \quad & P(x, c_k) = f(x) + c_k \sum_{i=1}^m \ln\left(\left(g_i^+(x)\right)^2 + 1^i\right), \\ & i \in I = \{1, 2, \dots, m\} \\ \text{subject to } & x \in R^n. \end{aligned} \tag{10}$$

We should call the problem (10) modified Courant-Beltrami penalized optimization problem.

Theorem 2 [13]. Suppose that f, g , and p are continuous functions. Let $\{x_k\}, k = 1, 2, \dots$ be a sequence of solutions in the penalized problem (10). Then any limit of the convergent subsequence $\{x_{k_j}\}$ solves the original nonlinear programming problem (1).

4 Karush-Kuhn-Tucker multiplier for the penalized problem

In optimization problems, the first order necessary conditions for a nonlinear optimization problem to be optimal is Karush-Kuhn-Tucker (KKT) conditions, or Kuhn-Tucker conditions if some of the constraint

qualifications are satisfied. Nevertheless, Courant-Beltrami penalty function may not be differentiable at a point $g_i(x) = 0$ for some $i \in I$. But for the constrained optimization both objective function and constraints may be partially differentiable on R^n while the penalized problem is not, since the differentiability is not among the properties of $\max\{0, g_i(x)\}$. According to Proposition 1 [21], some additional assumptions may be imposed on the constraint function $g_i(x)$, i.e. if the constraint $g_i(x)$ has continuous first-order partial derivatives on R^n , for this reason, $(g_i^+(x))^2$ admit the same. Therefore,

$$\frac{\partial}{\partial x_j} (g_i^+(x))^2 = 2(g_i^+(x)) \frac{\partial}{\partial x_j} g_i(x). \tag{11}$$

Considering equation (11), if $p(x): R^n \rightarrow R$ is a modified Courant-Beltrami (MCB) penalty function and the constraints $g_i(x)$ has continuous first-order partial derivative on R^n , then

$$\begin{aligned} \nabla p(x) &= \sum_{i=1}^m \nabla \left[\ln\left(\left(g_i^+(x)\right)^2 + 1^i\right) \right] \\ &= \sum_{i=1}^m \frac{2g_i^+(x) \nabla g_i(x)}{\left(\left(g_i^+(x)\right)^2 + 1^i\right)} \end{aligned}$$

Theorem 3. Let x^* be an optimal solution to the problem (1) and it satisfies the first-order necessary optimality conditions of the constrained problem (1). Moreover, a suitable CQ be satisfied at x^* , then x^* is a solution to the penalized problem (10).

Proof. If x^* is a feasible point which satisfies the first-order necessary optimality conditions of the problem, then

$$\nabla f(x^*) + c_k \nabla p(x^*) = 0,$$

That is

$$\nabla f(x^*) + c_k \sum_{i=1}^m \frac{2g_i^+(x^*) \nabla g_i(x^*)}{\left(\left(g_i^+(x^*)\right)^2 + 1^i\right)} = 0. \tag{12}$$

Let us define

$$\xi_i^* = c_k \frac{2g_i^+(x^*)}{\left(\left(g_i^+(x^*)\right)^2 + 1^i\right)}. \tag{*}$$

Then (12) can be rewritten as

$$\nabla f(x^*) + \sum_{i=1}^m \xi_i^* \nabla g_i(x^*) = 0,$$

where ξ_i^* is a vector of KKT multiplier and $\xi_i^* \geq 0$ for all $i \in I$. Therefore, Lagrange multipliers for the penalized problem (10) can be represented as $\xi^* = [\xi_1^*, \xi_2^*, \dots, \xi_m^*]$. □

5 Lagrangian duality

We consider the following dual problem for the considered nonlinear optimization problem (1)

$$\phi(\xi) = \inf_{\xi \geq 0} \{L(x, \xi) : x \in R^n\} \rightarrow \max \quad (13)$$

Definition 5. A vector ξ is said to be a feasible solution for the dual problem (13) if $\phi(\xi) = \inf_{\xi \geq 0} \{L(x, \xi) : x \in R^n\} > -\infty$.

Note that if the problem (13) has at least one feasible vector ξ then the problem is called a consistent and any feasible vector $\bar{\xi}$ satisfying $\phi(\bar{\xi}) = \sup_{\xi \geq 0} \phi(\xi)$ is an optimal solution to the problem (13).

The relationship between primal and dual problem for the invex problem (i.e. *if all the functions constituting the problem are invex on R^n w.r.t. the same function η*) is either zero or a positive real number. Moreover, if a feasible solution of the dual problem returned a primal objective value lower bound, then it is referred as the weak duality, while in the absence of duality gap (*when their difference is zero*), then it is called strong duality.

Let $P = \inf \{f(x) : g_i(x) \leq 0, i = 1, 2, \dots, m, x \in R^n\}$ and $D = \sup \{\phi(\xi) : \xi \in R^m, \xi \geq 0\}$ be the primal and the dual optimal respectively. The following inequality related both optima: $P \geq D$ (if the problem (1) and (13) are consistent).

Let G represent the duality gap between the two problems that is, $G = P - D \geq 0$.

Theorem 4. (Weak duality theorem): Let the functions constituting the problem (1) be continuously differentiable on R^n . Suppose that the problem (1) and (13) are consistent, then their optima are finite and $G \geq 0$.

Proof. Let problem (1) and (13) has any feasible solution x and ξ , respectively. The feasibility of x implies that $g_i(x) \leq 0, i = 1, 2, \dots, m$. If ξ is feasible to the problem (13). Then $\xi_i g_i(x) \leq 0, i = 1, 2, \dots, m$. So,

$$f(x) \geq f(x) + \sum_{i=1}^m \xi_i g_i(x). \quad (14)$$

The right-hand side of the inequality (14) is Lagrange function. Therefore, $f(x) \geq L(x, \xi)$.

Since both x and ξ are randomly chosen feasible points in (1) and (13) respectively, then

$$f(x) \geq \phi(\xi) = \inf \{L(y, \xi) : y \in R^n\}. \quad (15)$$

It follows directly from inequality (15) that

$$f(x) \geq \sup_{\xi \geq 0} \left\{ \inf_{y \in R^n} L(y, \xi) \right\} = D. \quad (16)$$

Consequently,

$$P = \inf_{x \in F} f(x) \geq \phi(\xi). \quad (17)$$

Therefore, (16) and (17) justified that P and D are finite if (1) and (13) are consistent and $P \geq D$. This verified the relationship between the two problems. i.e. $G \geq 0$. \square

Definition 6. A continuous function $f(x)$ that is defined on R^n is coercive if $\lim_{x \rightarrow \infty} f(x) = +\infty$.

That is, for any constant $M > 0 \exists R_M > 0$ such that $\|f(x)\| > M$ whenever $\|x\| > R_M$.

Theorem 5. (Strong duality theorem): Let the functions constituting the problem (1) be continuously differentiable on R^n . Further, assume that those functions in the problem (1) are invex with respect to the same function η , also an objective function in (1) is coercive. If an optimization problem (1) is consistent then its dual problem (13) is also consistent and their optimal solutions coincide, that is, $G = 0$.

Proof. To prove Theorem 5, we use the modified Courant-Beltrami penalty function method. As described in this approach, we need to construct an unconstrained optimization problem as follows:

$$\begin{aligned} \text{minimize } P(x, c_k) &= f(x) + c_k \sum_{i=1}^m \ln \left((g_i^+(x))^2 + 1^i \right), \\ i \in I &= \{1, 2, \dots, m\}. \end{aligned}$$

Suppose that at the k th iterations, x_k is an optimal solution to the penalized problem (10). Therefore, for every positive integer c_k , $P(x_k, c_k) = \min \{P(x, c_k) : x \in R^n\}$. Moreover, the sequence $\{x_k\}$ is bounded and of all its convergent subsequences have limits as described in the penalty function method (see Prop. 15 [13]), these limits are the solutions to the original problem (1). By Theorem 3, it follows that

$$\nabla f(x_k) + c_k \sum_{i=1}^m \frac{2g_i^+(x_k) \nabla g_i(x_k)}{\left((g_i^+(x_k))^2 + 1^i \right)} = 0.$$

Since x_k is an optimal solution of the problem (10), let assume that $\{x_{k_j}\}$ is a convergent subsequence of $\{x_k\}$, then the condition below is satisfied.

$$\nabla P(x_{k_j}, c_{k_j}) = \nabla f(x_{k_j}) + \sum_{i=1}^m \frac{2c_{k_j} g_i^+(x_{k_j}) \nabla g_i(x_{k_j})}{\left((g_i^+(x_{k_j}))^2 + 1^i \right)} = 0. \quad (18)$$

According to KKT multiplier defined in (*), the Lagrange multipliers are given by

$$(\bar{\xi}_i)_j = c_{k_j} \frac{2g_i^+(x_{k_j})}{\left((g_i^+(x_{k_j}))^2 + 1^i \right)}, i = 1, \dots, m, \forall j. \quad (**)$$

Accordingly, for each j ,

$$(\bar{\xi})_j = \left((\bar{\xi}_1)_j, (\bar{\xi}_2)_j, \dots, (\bar{\xi}_m)_j \right).$$

From (**) $(\bar{\xi})_j$ is non-negative. Therefore, for each j , Lagrange multipliers satisfy $(\bar{\xi})_j \geq 0$, then, by (18) and (**), we have $\nabla L(x_{k_j}, (\bar{\xi})_j) = 0$.

Since the functions constituting the problem (1) are invex functions on R^n with respect to the same function η . Thus, based on Definition 3, $\forall x \in R^n$, we have

$$f(x) - f(x_{k_j}) \geq \nabla f(x_{k_j})\eta(x, x_{k_j}) \tag{19}$$

$$g(x) - g(x_{k_j}) \geq \nabla g(x_{k_j})\eta(x, x_{k_j}).$$

Then, multiplying the second part of (19) by the Lagrange multiplier vector, we get

$$(\bar{\xi})_j g(x) - (\bar{\xi})_j g(x_{k_j}) \geq (\bar{\xi})_j \nabla g(x_{k_j})\eta(x, x_{k_j}). \tag{20}$$

Upon combining and rearranging the first part of (19) and (20), we have

$$\begin{aligned} & [f(x) + (\bar{\xi})_j g(x)] - [f(x_{k_j}) + (\bar{\xi})_j g(x_{k_j})] \\ & \geq [\nabla f(x_{k_j}) + (\bar{\xi})_j \nabla g(x_{k_j})]\eta(x, x_{k_j}). \end{aligned} \tag{21}$$

Then, according to (21) the Lagrange function $L(x_{k_j}, (\bar{\xi})_j)$ is an invex function with respect to the same function η as in Definition 3 with stationary points $(x_{k_j}, (\bar{\xi})_j)$ according to equation $\nabla L(x_{k_j}, (\bar{\xi})_j) = 0$, and the points are global minimum of the function L . Therefore,

$$L(x_{k_j}, (\bar{\xi})_j) = \min\{L(x, (\bar{\xi})_j) : x \in R^n\} > -\infty. \tag{22}$$

From (22), it follows that the Lagrange multiplier vector $(\bar{\xi})_j$ is feasible for the dual problem (13). By hypothesis, the optimization problem (1) is consistent with the coercive objective function. Then, the problem has a solution with primal optimal value $P > -\infty$. Consequently, for each j

$$P(x_{k_j}, c_{k_j}) \geq f(x_{k_j}). \tag{23}$$

The reason for (23) is based on the assumptions that the limit point \bar{x} of the convergent subsequence $\{x_{k_j}\}$ is an optimal solution to the original problem (1). By the construction of modified Courant-Beltrami penalized problem, we have

$$P(x_{k_j}, c_{k_j}) = f(x_{k_j}) + c_{k_j} \sum_{i=1}^m \ln\left(\left(g_i^+(x_{k_j})\right)^2 + 1^i\right), \tag{24}$$

and considering the definition of $g^+(x)$, it is obvious that

$$P(x_{k_j}, c_{k_j}) \leq f(x_{k_j}) + \frac{\sum_{i=1}^m 2c_{k_j} \left(g_i^+(x_{k_j})\right)^2}{\left(\left(g_i^+(x_{k_j})\right)^2 + 1^i\right)}. \tag{25}$$

But x_{k_j} is not feasible in the original optimization problem, then in (25) $\left(g_i^+(x_{k_j})\right)^2$ can be expanded and rewritten in the following form:

$$P(x_{k_j}, c_{k_j}) \leq f(x_{k_j}) + \frac{\sum_{i=1}^m 2c_{k_j} g_i^+(x_{k_j}) g_i(x_{k_j})}{\left(\left(g_i^+(x_{k_j})\right)^2 + 1^i\right)} \tag{26}$$

Then, by (**) inequality (26) can be replaced with

$$P(x_{k_j}, c_{k_j}) \leq f(x_{k_j}) + \sum_{i=1}^m (\bar{\xi})_j g_i(x_{k_j}). \tag{27}$$

Therefore, according to definition of Lagrange function, we have

$$P(x_{k_j}, c_{k_j}) \leq L(x_{k_j}, (\bar{\xi})_j).$$

Since x_{k_j} is a convergent subsequence of $\{x_k\}$, then by (22)

$$L(x_{k_j}, (\bar{\xi})_j) = \min\{L(x, (\bar{\xi})_j) : x \in R^n\}.$$

So,

$$\min\{L(x, (\bar{\xi})_j) : x \in R^n\} \leq D. \tag{28}$$

Consequently, by (23) and (28)

$$f(x_{k_j}) \leq D \forall j. \tag{29}$$

Since the sequence $\{x_k\}$ converge to \bar{x} , which is optimal to the original problem (1) and f is a continuous function on R^n , then we conclude that $P = f(\bar{x}) = \lim_{j \rightarrow \infty} f(x_{k_j}) \leq D \Rightarrow P \leq D$.

Also, by Theorem 4, $P \geq D$. Combining the two results (i.e. $G \leq 0$ and $G \geq 0$), $P = D$. \square

The results established in this paper will be illustrated with some examples, using a modified Courant-Beltrami penalty function method. Precisely, for invex optimization problem.

Example 1. [22] Consider the following optimization problem:

$$\min f(x) = \ln(x_1 + 1) + (x_2^2 + 1) \arctan x_2$$

$$g_1(x) = -\ln(x_1 + 1) \leq 0$$

$$g_2(x) = -(x_2^2 + 1) \arctan x_2 \leq 0$$

$$X = \{(x_1, x_2) \in R^2 : x \geq -1\}$$

This problem can be verified that the objective function f and both constraints functions g_1 and g_2 are invex functions on the set X with respect to the same function η defined by

$$\eta(x, u) = \begin{pmatrix} \eta_1(x, u) \\ \eta_2(x, u) \end{pmatrix}$$

where

$$\eta_1(x, u) = (u_1 + 1)[\ln(x_1 + 1) - \ln(u_1 + 1)],$$

$$\eta_2(x, u) = \frac{1}{1 + 2u_2 \arctan u_2} [(x_2^2 + 1) \arctan x_2 - (u_2^2 + 1) \arctan u_2].$$

Modified Courant-Beltrami penalty function method defined by (10) can be used to solve the problem. Converting the problem in to unconstrained problem, we obtained the following:

$$\begin{aligned} P(x, c) = & \ln(x_1 + 1) + (x_2^2 + 1) \arctan x_2 \\ & + c \left[\ln((\max\{-\ln(x_1 + 1), 0\})^2 + 1) \right. \\ & \left. + c \ln\left(\left(\max(-(x_2^2 + 1) \arctan x_2, 0)\right)^2 + 1\right) \right]. \end{aligned} \quad (30)$$

Note that, the set of feasible solutions $F = \{(x_1, x_2) \in X : x_1 \geq 0, x_2 \geq 0\}$. The points $(0, 0)$ is an optimal to the problem, since the conditions (2)–(4) are satisfied with Lagrange multipliers $\bar{\xi}_1 = \bar{\xi}_2 = 1$, and penalty parameter $c > \max\{|\bar{\xi}_i|, i = 1, 2\} = 1$, it is also a minimizer to penalized problem (30).

Furthermore, according to the definition of Lagrange function, the Lagrangian dual for the considered problem can be constructed as follows:

$$\begin{aligned} \phi(\xi) = & \min\{\ln(x_1 + 1) + (x_2^2 + 1) \arctan x_2 + \xi_1 \\ & (-\ln(x_1 + 1)) + \xi_2(-(x_2^2 + 1) \arctan x_2)\} \rightarrow \max \end{aligned}$$

The optimal values of $\phi(\xi)$ and $P(x, c)$ are the same considering that all the stated assumptions in Theorem 5 are satisfied, the duality gap is 0 (i.e. $G = 0$).

Let us now consider the case when some assumptions in theorem were not fulfilled. In this case, the duality gap may no longer be equal to 0 ($G \neq 0$).

Example 2. [13] Consider the following optimization problem:

$$\min f(x) = x^4 - \frac{4}{3}x^3 - 4x^2$$

$$g(x) = x - \frac{1}{2} \leq 0.$$

Obviously, the objective f is coercive, but it is not invex on R with respect to any function $\eta : R \times R \rightarrow R$, due to the facts that the stationary point $x = 0$ of f is not its global minimum point.

The penalized problem can be constructed as follows:

$$\begin{aligned} P(x, c) = & x^4 - \frac{4}{3}x^3 - 4x^2 \\ & + c \ln\left(\left(\max\left\{x - \frac{1}{2}, 0\right\}\right)^2 + 1\right). \end{aligned} \quad (31)$$

The feasible point $x = -1$ is an optimal in (31) with the optimal value $P(-1, c) = \frac{-5}{3}$. Thus, the Lagrangian dual as defined by (13) is as follows:

$$\begin{aligned} \phi(\xi) = & \inf_{x \in R} L(x, \xi) \\ = & \inf_{x \in R} \left\{ x^4 - \frac{4}{3}x^3 - 4x^2 + \xi \left(x - \frac{1}{2} \right) \right\}. \end{aligned} \quad (32)$$

Clearly, the dual problem (32) is consistent since $\inf_{x \in R} L(x, \xi) > -\infty$ with $\sup \phi(\xi) = \frac{-37}{6}$. Hence $G = \frac{9}{2}$ which is greater than 0.

6 Conclusion

In this paper, we modified a penalty function that should be called a modified Courant-Beltrami (MCB) penalty function method. We combined the concept of Lagrangian dual and penalty function approach to investigate and prove the relationship between an optimization problem (1) and its Lagrangian dual problem (13). The zero-duality gap for invex optimization problem has been established under some moderate assumptions. Some examples were presented which verified that those assumptions are essential to establish a zero-duality gap between the two problems. In the future research, MCB penalty function method will be applied to solve;

- Multi-objective constrained optimization problem and
- The practical problem from any of the following areas;
 - Engineering
 - Decision theory and
 - The chemical process, etc.

References

1. W.I. Zangwill, Nonlinear programming via penalty functions, *Manag. Sci.* **13**, 344 (1967)
2. B.W. Kort, D.P. Bertsekas, Combined primal-dual and penalty method for convex programming, *SIAM J. Cont. Optim.* **14**, 268 (1976)
3. R.V. Rao, Jaya: a simple and new optimization algorithm for solving constrained and unconstrained optimization problems, *Int. J. Ind. Eng. Comput.* **7**, 19 (2016)
4. A. Cherukuri, E. Mallada, J. Cortes, Asymptotic convergence of constrained primal-dual dynamics, *Syst. Control Lett.* **87**, 10 (2016)
5. C. Kanzow, D. Steck, Augmented Lagrangian and exact penalty methods for quasi-variational inequalities. *Comput. Optim. Appl.* **69**, 801 (2017)
6. R.A. Shandiz, E. Tohidi, Decrease of the penalty parameter in differentiable penalty function methods, *Theor. Econ. Lett.* **1**, 8 (2011)
7. G.D. Pillo, L. Grippo, A continuously differentiable exact penalty function for nonlinear programming problems with inequality constraints, *SIAM J. Cont. Optim.* **23**, 72 (1985)

8. Z. Chen, Y. Dai, A line search exact penalty method with bi-object strategy for nonlinear constrained optimization, *J. Comput. Appl. Math.* **300**, 245 (2016)
9. X.L. Sun, D. Li, Logarithmic-exponential penalty formulation for integer programming, *Appl. Math. Lett.* **12**, 73 (1999)
10. M. Bazaraa, H. Sherali, C. Shetty, *Nonlinear Programming Theory and Algorithms* (John Wiley & Sons, Inc. New Jersey, 2006), 3rd edn.
11. W. Changyu, X. Yang, X. Yang, Nonlinear lagrange duality theorems and penalty function methods in continuous optimization, *J. Glob. Optim.* **27**, 473 (2003)
12. D.Y. Gao, N. Ruan, H.D. Sherali, Solutions and optimality criteria for nonconvex constrained global optimization problems with connections between canonical and Lagrangian duality, *J. Glob. Optim.* **45**, 473 (2009)
13. T. Antczak, Penalty function methods and a duality gap for invex optimization problems, *Nonlinear Anal. Theory Methods Appl.* **71**, 3322 (2009)
14. M. Hanson, On sufficiency of the Kuhn-Tucker conditions in nondifferentiable programming, *J. Math. Anal. Appl.* **80**, 545 (1981)
15. B.D. Craven, Invex functions and constrained local minima, *Bull. Aust. Math. Soc.* **24**, 357 (1981)
16. E. Ernst, M. Volle, Generalized Courant-Beltrami penalty functions and zero duality gap for conic convex programs, *Positivity* **17**, 945 (2013)
17. N. Kanzi, M. Soleimani-damaneh, Slater CQ, optimality and duality for quasiconvex semi-infinite optimization problems, *J. Math. Anal. Appl.* **434**, 638 (2016)
18. M. Soleimani-damaneh, Penalization for variational inequalities, *Appl. Math. Lett.* **22**, 347 (2009)
19. M. Soleimani-damaneh, Nonsmooth optimization using Mordukhovich's subdifferential, *SIAM J. Control Optim.* **48**, 3403 (2010)
20. M. Hassan, A. Baharum, A new logarithmic penalty function approach for nonlinear constrained optimization problem, *Deci. Sci. Lett.* **8**, 3 (2019)
21. D.P. Bertsekas, On penalty and multiplier methods for constrained minimization, *SIAM J. Cont.* **14**, 216 (1976)
22. T. Antczak, Exact penalty functions method for mathematical programming problems involving invex functions, *Eur. J. Oper. Res.* **198**, 29 (2009)

Cite this article as: Mansur Hassan, Adam Baharum, Modified courant-Beltrami penalty function and a duality gap for invex optimization problem, *Int. J. Simul. Multidisci. Des. Optim.* **10**, A10 (2019)