Generalized logarithmic penalty function method for solving smooth nonlinear programming involving invex functions

Mansur Hassan\textsuperscript{a,b} and Adam Baharum\textsuperscript{a} 

\textsuperscript{a}School of Mathematical Sciences, Universiti Sains Malaysia, Geligor Pulau Pinang, Malaysia; \textsuperscript{b}Department of Mathematical Sciences, Yusuf Maitama Sule University Kano, Kano, Nigeria

ABSTRACT

In this paper, we have reviewed some penalty function methods for solving constrained optimization problems in the literature and proposed a continuously differentiable logarithmic penalty function which consists of the proposed logarithmic penalty function and modified Courant-Beltrami penalty function for equality and inequality constraints, respectively. Furthermore, we hybridized the two and came up with the general form of both (equality and inequality) constraints. However, in the first part, the equivalence between the sets of optimal solutions in the original optimization problem and its associated penalized logarithmic optimization problem constituted by invex functions with equality and inequality constraints has been established. In the second part, we have validated the general form of the logarithmic penalty function and compared the results with absolute value penalty function results by solving nine small problems from Hock-Schittkowski collections of test problems with different classifications. The experiments were carried out via quasi-newton algorithm using a \texttt{fminunc} routine function in \texttt{matlab2018a}. The general form yields a better objective value compared to absolute value penalty function.

1. Introduction

In this paper, the general form of constrained optimization involving equality and inequality constraints is considered:

\begin{equation}
\text{Minimize } f(x) \\
\text{Subject to } h_j(x) = 0, \ j \in J = \{1, 2, \ldots, s\}, \ g_i(x) \leq 0, \ i \in I = \{1, 2, \ldots, m\}, \ x \in X,
\end{equation}

Where $f: X \rightarrow \mathbb{R}$ and $h_j: X \rightarrow \mathbb{R}$, $j \in J$. $g_i: X \rightarrow \mathbb{R}$, $i \in I$, are differentiable functions on a non-empty open subset of real number $X$. Let’s introduce some notation that will frequently be used throughout this paper.

Let $\mathcal{X} = \{x \in X : h_j(x) = 0, \ j \in J, \ g_i(x) \leq 0, \ i \in I\}$ be the set of all feasible solutions of the optimization problem (P).

$I(X) = \{i \in I : g_i(X) = 0\}$ denotes the set of active constraints at an optimal point $X$.

In recent years, an approach to solve a non-linear optimization problem (P) is not limited to the conventional strategies for solving a constrained optimization problem only. Several methods have come into existence in the last few decades. One of the attractive approaches is the use of the penalty function, and it is usually defined in terms of the constraints $h$ and $g$. The constraints are incorporated into the objective function by adding a penalty term that penalized any violations of the single unconstrained optimization problem. This approach guaranteed that the solution could be found through unconstrained minimization technique.

Based on the existing approach of penalty function methods, it has been divided into two forms: a standard penalty function and an exact penalty function which can also be subdivided into differentiable and non-differentiable.

Zangwill (1967) was the first to introduce non-differentiable exact penalty function and presented an algorithm that can be used to solve a non-linear programming problem, and the method appears to be more useful in the concave case. Primarily, a penalty function has been used to transform a problem (P) into a single unconstrained problem or a finite sequence of the unconstrained optimization problems. The non-differentiable exact penalty function introduced by Zangwill (1967) for the problem (P) was:

\begin{equation}
p(x) = \sum_{i=1}^{m} g_i^p(x) + \sum_{j=1}^{s} |h_j(x)|
\end{equation}
Note that \( g_i^+(x) = \max \{0, g_i(x)\} \).

The work of Fletcher (1973) comes up with a continuously differentiable exact penalty function regarding the problem (P) for inequality constraints; according to Fletcher’s work it is possible to construct an exact penalty function which is sufficiently smooth to grant conventional techniques for solving the problem (P) to locate the local minimum for the transformed unconstrained problem. Other researchers (Bazaraa, Sherali, & Shetty, 1991; Bertsekas & Koksal, 2000; Charalambous, 1978; Conn, 1973) further investigated continuously differentiable and non-differentiable exact penalty function.

However, the idea of convexity plays a dominant role in almost all kinds of penalty function approaches (see for example Bazaraa et al. (1991), Bertsekas and Koksal (2000), Charalambous (1978), Conn (1973) and Mangasarian (1985)). In the last few years, some various convex function generalizations have been derived which give room for extending optimality condition and some classical duality results, formerly restricted to convex programmes, to the larger classes of optimization problems. The notion of invexity introduced by Hanson (1981) and named by Craven (1981) was among the types. The work of Hanson applied the extended concept of convex functions to prove optimality conditions and duality results for the non-linear constrained optimization problem.

Antczak (2009) presented some new results on the exact penalty function methods; the work of Antczak characterized a differentiable non-convex optimization problem with mixed constraints as in problem (P) via the following exact penalty function with \( q = 1 \):

\[
p(x) = \sum_{i=1}^{m} [g_i^+(x)]^q + \sum_{j=1}^{l} |h_j(x)|^q\tag{2}
\]

where \( q \) is a positive integer, the equivalence between sets of optimal solutions in the problem (P) and the following penalized optimization problem under suitable invexity assumption is established:

\[
P(x,c) = f(x) + c \left[ \sum_{i=1}^{m} [g_i^+(x)] + \sum_{j=1}^{l} |h_j(x)| \right] \tag{3}
\]

where \( c \) is a penalty parameter.

Liu and Feng (2010) constructed a classical exponential penalty function method for multi-objective programming problems (MOPP) and its convergence is proved. Antczak (2011) introduced the \( l_1 \) exact exponential penalty function method, which is constructed based on the optimization problem (P) with both constraints (equality and inequality). The \( l_1 \) exact exponential penalty function method is used to solve an optimization problem constituted by \( r \)-invex functions (with respect to the same function \( \eta \), the penalty function is of the following form:

\[
p(x) = \sum_{i=1}^{m} \frac{1}{r} (e^{g_i^+(x)} - 1) + \sum_{j=1}^{l} \frac{1}{r} (e^{|h_j(x)|} - 1), \tag{4}
\]

where \( r \) is a finite real number not equal to 0. Note that the function \( \frac{1}{r} (e^{g_i^+(x)} - 1) \) is defined by

\[
\frac{1}{r} (e^{g_i^+(x)} - 1) = \begin{cases} 
0, & \text{if } g_i(x) \leq 0, \\
\frac{1}{r} (e^{g_i^+(x)} - 1), & \text{if } g_i(x) > 0
\end{cases}
\]

Obviously, (5) has the penalty features relative to the single constraint \( g_i(x) \leq 0 \), that is 0 for all values of \( x \) that satisfy the constraint and result in large values whenever the constraint is violated; the penalty function (4) is considered to be a classical (1) if \( r = 0 \) that was defined by Pietrzykowski (1969) and also by Han and Mangasarian (1979). Further, the results have been proved through the classical \( l_1 \) exact penalty function method under \( r \)-invexity assumption by Antczak (2010) for both equality and inequality constraints. Jayswal and Choudhury (2014) were able to extend the work of Antczak (2011) and Liu and Feng (2010) to multiobjective fractional programming problems and examine the convergence of the method. Morrison (1968) proposes another penalized problem of the form:

\[
\min |f(x) - M|^2 + |h(x)|^2 \tag{6}
\]

where \( M \) is an estimated optimized objective function \( f(x^*) \) and \( h(x) \) is an equality constraint. The Morrison function (6) can also be considered as an exact penalty function. Another, similar to that of the Morrison method has been proposed by Meng, Hu, Dang, and Yang (2004) and Meng, Dang, Jiang, Xu, and Shen (2013). Luenberger (1973) explored an unconstrained problem that works in the space \( \mathbb{R}^{m+n} \) with respect to the objective and constraint functions in (6) of the following form:

\[
\min |\nabla f(x) - \lambda^T \nabla h(x)|^2 + |h(x)|^2 \tag{7}
\]

where \( \lambda^T \) is the transpose of LaGrange multiplier vector; this approach does not require successive minimization solution. Nevertheless, this approach admits disadvantage of higher dimension.

Filter based approach to solve the same constrained optimization problem were introduced by Fletcher and Leyffer (2002). The concept is implemented by minimizing two functions \( f(x), \Gamma(x) \) simultaneously. The function \( \Gamma(x) \) possesses the basic penalty function property, i.e.

(i) \( \Gamma(x) > 0 \), if \( x \) is infeasible

(ii) \( \Gamma(x) = 0 \), if \( x \) is feasible

It is a list of pairs \( (f_i, \Gamma_i) \) whereas no pair will be allowed to dominate another. Nie (2007) modified the original filter method specifically for the equality constraints.
constraint only, and the process is called a sequential penalty quadratic programming filter method (SIQP). His approach replaced the objective function by the penalized function of the form:

\[ f(x) + \sigma \Gamma(x) \]  

where \( \sigma \) is a fixed parameter that does not need to be updated at each step. According to Nie, this approach is advantageous compared to the original filter method.

Utsch De Freitas Pinto and Martins Ferreira (2014) propose an exact penalty function based on matrix projection, and the constructed unconstrained problem is of the form:

\[ \min |d(x)|^2 + |h(x)|^2 \]  

where \( d(x) = P(x)^T f(x) \) (gradient projection vector used by Rosen (1960, 1961)) \( P(x)^T f(x) \) is the transpose of \( \nabla f(x) \) and \( P(x) = I - \nabla^T h(x) \left( \nabla h(x)^T \nabla h(x) \right)^{-1} \nabla h(x) \) (projection matrix over the constrained tangent subspace of the considered problem).

Some of the observed setbacks of this approach include; all the stationary points of the considered problem are local minima and the difficulty of the matrix inversion required to compute \( P(x) \). Moreover, none of the above-listed penalty functions approach the authors’ claim to be perfect on any formulations. Another critical point to be taken into consideration is that the composition of any optimization problem may vary from one form to another, apart from the notion of convex, invex, non-convex optimization problems; there are also regularity, irregularity, linear objective with non-linear constraints functions, non-linear objective with linear constraints functions, quadratic and polynomial functions.

In this paper, we propose another penalty function in accordance with existing penalty function methods. The proposed penalty function is constructed according to the general form of mathematical programming problems (P), since it is suitable for the optimization problems with mixed constraints (equality and inequality), unlike some of the existing ones that are either restricted to equality constraints or inequality constraints only. Furthermore, it is among the category of the continuously differentiable penalty function. The entire presentation is organized as follows: In Section 2, the method of the proposed logarithmic penalty function is introduced. In Section 3, we use the first order necessary optimality conditions to derive the KKT multipliers for the proposed logarithmic penalty function. In Section 4, the equivalence between the sets of optimal solutions of the original optimization problem and its associated penalized problem constituted by invex functions (with respect to the same function \( \eta \)) was established. In Section 5, we validated the proposed general form of logarithmic penalty function methods via a quasi-Newton algorithm using the \texttt{fminunc} routine function in matlab2018a. Finally, in the last section, the results obtained are discussed followed by the conclusion of the paper.

2. Method of logarithmic penalty function

Transforming a constrained problem into a single or a sequence of unconstrained optimization problems can be done through penalty methods. The primary responsibility of the penalty term is to incur a decisive penalty for infeasible points, while there is no penalty for feasible points.

For the considered non-linear optimization problem (P), its penalized optimization problem is defined by \( P(x, c) = f(x) + cp(x) \).

\( p \) represents a suitable penalty function, and \( c \) is a positive penalty parameter. Now, the newly constructed logarithmic penalty function consists of the following parts:

(i) Proposed logarithmic penalty function for equality constraints

(ii) Modified Courant-Beltrami penalty function for inequality constraints

If we consider an optimization problem (P) with equality constraints only, we construct the following penalty function regarding (i):

\[ p_1(x) = \sum_{j=1}^{s} \ln \left[ \left( h_j(x) \right)^2 + 1 \right] \]  

It is obvious that (10) possesses the penalty function features relative to the equality constraints \( h_j(x) = 0 \) as presented by Hassan and Baharum (2019); the penalized optimization based on (10) should be constructed as follows:

\[ P(x, c) = f(x) + c \sum_{j=1}^{s} \ln \left[ \left( h_j(x) \right)^2 + 1 \right] \]  

To make the proposed logarithmic penalty function more general, we modified the Courant-Beltrami penalty function as presented in (Byrne, 2008; Ernst & Volle, 2013), and the Courant-Beltrami penalty function restricted to inequality constraints is constructed based on the exact penalty function (2) with \( q = 2 \) as follows:

\[ p(x) = \sum_{i=1}^{m} \left[ g_i^+(x) \right]^2 \]  

Therefore, we modified (12) in logarithmic form like the proposed penalty function (10) as follows:

\[ p_2(x) = \sum_{i=1}^{m} \ln \left[ \left( g_i^+(x) \right)^2 + 1 \right] \]
Combining (10) and (13) will lead to generalized logarithmic penalty function regarding the problem (P), i.e.

\[ p(x) = p_1(x) + p_2(x) \]  

(14)

For simplicity and convenience, upon substitution of (11) and (13) right hand sides into (14), the generalized logarithmic penalty function can be expressed as follows:

\[ p(x) = \sum_{j=1}^{s} \ln \left( (h_j(x))^2 + 1 \right) + \sum_{i=1}^{m} \ln \left( (g_i^*(x))^2 + 1 \right) \]

(15)

These terms can also be denoted by

\[ \varphi(h_j(x)) = \ln \left( (h_j(x))^2 + 1 \right) \]

\[ \phi(g_i(x)) = \ln \left( (g_i^*(x))^2 + 1 \right) \]

\[ p(x) \] can simply be rewritten as follows:

\[ p(x) = \sum_{j=1}^{s} \varphi(h_j(x)) + \sum_{i=1}^{m} \phi(g_i(x)) \]

Both functions \( \varphi \) and \( \phi \) are continuous functions satisfying the following conditions:

(i) \( \varphi(h_j(x)) = 0 \) if \( h_j(x) = 0 \) and \( \varphi(h_j(x)) > 0 \) if \( h_j(x) \neq 0 \) for \( j = 1, 2, \ldots, s \)

(ii) \( \phi(g_i(x)) = 0 \) if \( g_i(x) \leq 0 \) and \( \phi(g_i(x)) > 0 \) if \( g_i(x) > 0 \) for \( i = 1, 2, \ldots, m \)

3. Logarithmic penalty function Karush-Kuhn-Tucker multipliers

In optimization problems, the first order necessary conditions for a non-linear optimization problem to be optimal are Karush-Kuhn-Tucker (KKT) conditions, or Kuhn-Tucker conditions, if some of the constraints’ 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, as differentiability is not among the properties of \( \max \{0, g_i(x)\} \). According to proposition 1 (Ernst & Volle, 2013), 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,

\[ \frac{\partial}{\partial x_l} (g_i^*(x))^2 = 2(g_i^*(x)) \frac{\partial}{\partial x_l} g_i(x) \]

(*)

where \( l = 1, 2, \ldots, n \), representing the number of variables.

Considering equation (*), if \( p_2(x) : R^n \rightarrow R \) is a modified Courant-Beltrami penalty function (MCBP) and the constraints \( g_i(x) \) have continuous first order partial derivative on \( R^n \), then

\[ \nabla p_2(x) = \sum_{i=1}^{m} \nabla \left[ \ln \left( (g_i^*(x))^2 + 1 \right) \right] \]

\[ = \sum_{i=1}^{m} \left( \frac{2g_i^*(x) \nabla g_i(x)}{(g_i^*(x))^2 + 1} \right) \]

(***)

In the same way, for equality constrained let \( p_1(x) : R^n \rightarrow R \), and assume that \( p(x) \) is continuously differentiable at \( x \), then

\[ \nabla p_1(x) = \sum_{j=1}^{s} \nabla \left[ \ln \left( (h_j(x))^2 + 1 \right) \right] \]

\[ = \sum_{j=1}^{s} \left( \frac{2h_j(x) \nabla h_j(x)}{(h_j(x))^2 + 1} \right) \]

(****)

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

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

\[ \nabla f(x^*) + c_i \nabla p(x^*) = 0, \]

From (**) and (***) we have

\[ \nabla f(x^*) + c_i \left( \sum_{j=1}^{s} \frac{2h_j(x) \nabla h_j(x)}{(h_j(x))^2 + 1} + \sum_{i=1}^{m} \frac{2g_i^*(x) \nabla g_i(x)}{(g_i^*(x))^2 + 1} \right) = 0 \]

(16)

Let us define

\[ \mu^*_i = c_k \left( \frac{2h_j(x^*)}{(h_j(x^*))^2 + 1} \right) \]

\[ \xi^*_i = c_k \left( \frac{2g_i^*(x^*)}{(g_i^*(x^*))^2 + 1} \right) \]

Then (16) can be replaced with the following equation;

\[ \nabla f(x^*) + \sum_{i=1}^{s} \mu^*_i \nabla h_i(x^*) + \sum_{i=1}^{m} \xi^*_i \nabla g_i(x^*) = 0 \]

where \( \mu^*_i \) and \( \xi^*_i \) are vectors of KKT multiplier and \( \mu^*_i \geq 0 \) for all \( i \in I \), \( \mu^*_i \geq 0 \) for all \( j \in J \).

Theorem 2. Let \( x^* \) be an optimal solution to the problem (P) and assume that any constraint qualification is satisfied at \( x^* \). Then there exist LaGrange multipliers \( \mu^*_i \in R^i \) and \( \xi^*_i \in R^m \) such that

\[ \nabla f(x^*) + \sum_{i=1}^{s} \mu^*_i \nabla h_i(x^*) + \sum_{i=1}^{m} \xi^*_i \nabla g_i(x^*) = 0 \]  

(i)

\[ \xi^*_i g_i(x^*) = 0, i = 1, 2, \ldots, m \]  

(ii)

\[ \xi^*_i \geq 0 i = 1, 2, \ldots, m \]  

(iii)

Definition 1. A point \( x^* \in F \) is said to be a KKT point if the condition (i)-(iii) is satisfied at \( x^* \).
Definition 2. (Hanson, 1981) Let \( f : X \to R \) be a differentiable function on \( X \subset R^n \) and \( u \in R^n \). If there exists a vector-valued function \( \eta : R^n \times R^n \to R^n \) such that, \( \forall \, x \in X \), the following inequality

\[
f(x) - f(u) \geq \nabla f(u) \cdot \eta(x,u) (> 0)
\]

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

If (17) holds at each point \( u \in R^n \), then \( f \) is said to be an invex (strictly invex) function with respect to \( \eta \) on \( R^n \).

Definition 3. (Hanson, 1981) Let \( f : X \to R \) be a differentiable function on \( X \subset R^n \) and \( u \in R^n \). If there exists a vector-valued function \( \eta : R^n \times R^n \to R^n \) such that, \( \forall \, x \in X \), the following inequality

\[
f(x) - f(u) \leq \nabla f(u) \cdot \eta(x,u) (< 0)
\]

holds, then the function \( f \) is said to be an invcave (strictly invcave) function with respect to \( \eta \) at \( u \) on \( X \).

If (18) holds at each point \( u \in R^n \), then \( f \) is said to be an invcave (strictly invcave) function with respect to \( \eta \) on \( R^n \).

4. Logarithmic penalty function method for invex optimization problem

In this section, we consider the case whereby all the functions constituting an optimization problem are invex (with respect to the same function \( \eta \)). Under invexity assumption imposed on the functions constituting the problem (P) and the relevant constraints qualifications (Bazaraa et al., 1991), for a suitably large penalty parameter \( c \) exceeding some suited Threshold, a KKT point minimizes the problem (P) if and only if it minimizes the problem defined by:

\[
P(x,c) = f(x) + c \left( \sum_{j=1}^{s} \ln \left( (h_j(x))^2 + 1 \right) + \sum_{i=1}^{m} \ln \left( (g_i^+(x))^2 + 1 \right) \right)
\]

or simply

\[
P(x,c) = f(x) + c \left( \sum_{j=1}^{s} \phi(h_j(x)) + \sum_{i=1}^{m} \phi(g_i(x)) \right)
\]

The subsequent theorem establishes the equivalence between an optimal solution in the original optimization problem (P) and its logarithmic penalty function minimizer under assumptions that the problem (P) is an invex optimization problem.

Theorem 3. Let \( x^* \) be a feasible point in the considered optimization problem (P) and KKT necessary optimality conditions (i)–(iii) fulfilled at \( x^* \). Partitioning the set of indexes for equality constraints \( h_j \) into \( J^+ \) and \( J^- \) i.e. \( J^+ = \{ j \in J : \mu_i^j > 0 \} \) and \( J^- = \{ j \in J : \mu_i^j < 0 \} \). Moreover, suppose that the function \( f, g_i, i \in I(x^*) \), \( h_j, j \in J^+ \) is invex at \( x^* \) on \( X \) with respect to the same function \( \eta \) and the function \( h_j, j \in J^- \) are invcave at \( x^* \) on \( X \) with respect to the same function \( \eta \). If \( c \) is sufficiently large (setting \( c \geq \max \{ \xi_i^j, \ \xi_i^j, \ i \in I, \ |\mu_i^j|, \ j \in J \} \).

Proof.

By hypothesis, \( x^* \) is feasible in the problem (P) and KKT conditions when (i)–(iii) are satisfied at the feasible point \( x^* \) with the Lagrange multiplier \( \mu_i^j \), \( j \in J, \ \xi_i^j, \ i \in I \), for equality constraints and inequality constraints, respectively.

Our goal is to show that \( x^* \) is also a minimizer of the penalized problem (19) with a sufficiently large penalty parameter \( c \). According to the problem (19), we have

\[
P(x,c) = f(x) + c \left( \sum_{j=1}^{s} \ln \left( (h_j(x))^2 + 1 \right) + \sum_{i=1}^{m} \ln \left( (g_i^+(x))^2 + 1 \right) \right)
\]

We set \( c \geq \max \{ \xi_i^j, \ i \in I, \ |\mu_i^j|, \ j \in J \} \), then

\[
f(x) + c \sum_{j=1}^{s} \ln \left( (h_j(x))^2 + 1 \right) + c \sum_{i=1}^{m} \ln \left( (g_i^+(x))^2 + 1 \right) \\
\geq f(x) + \sum_{j=1}^{s} \mu_i^j |h_j(x)| + \sum_{i=1}^{m} \xi_i^j g_i^+(x)
\]

Moreover,

\[
f(x) + \sum_{j=1}^{s} \mu_i^j |h_j(x)| + \sum_{i=1}^{m} \xi_i^j g_i^+(x) \\
\geq f(x) + \sum_{j=1}^{s} \mu_i^j h_j(x) + \sum_{i=1}^{m} \xi_i^j g_i^+(x)
\]

On the other hand

\[
f(x) + \sum_{j=1}^{s} |\mu_i^j h_j(x)| + \sum_{i=1}^{m} \xi_i^j g_i(x) \\
\geq f(x) + \sum_{j=1}^{s} \mu_i^j h_j(x) + \sum_{i=1}^{m} \xi_i^j g_i(x) \\
f(x) + \sum_{j=1}^{s} |\mu_i^j h_j(x)| + \sum_{i=1}^{m} \xi_i^j g_i(x) \\
\geq f(x) + \sum_{j=1}^{s} \mu_i^j [h_j(x^*) + \nabla h_j(x^*) \eta(x,x^*)] + \sum_{i=1}^{m} \xi_i^j [g_i(x^*) + \nabla g_i(x^*) \eta(x,x^*)] \\
+ \sum_{i=1}^{m} \xi_i^j [g_i(x^*) + \nabla g_i(x^*) \eta(x,x^*)] \\
\geq f(x) + \sum_{j=1}^{s} \mu_i^j [h_j(x^*) + \nabla h_j(x^*) \eta(x,x^*)] + \sum_{i=1}^{m} \xi_i^j [g_i(x^*) + \nabla g_i(x^*) \eta(x,x^*)]
\]
From KKT condition (ii), we have

\[
f(x) + \sum_{j=1}^{s} \mu_j^+ \langle h_j(x^*) + \nabla h_j(x^*) \eta(x^*) \rangle + \sum_{j=1}^{m} \zeta^+_i \langle g_i(x^*) + \nabla g_i(x^*) \eta(x^*) \rangle = f(x) + \sum_{j=1}^{s} \mu_j^+ \nabla h_j(x^*) \eta(x^*) + \sum_{j=1}^{m} \zeta^+_i \nabla g_i(x^*) \eta(x^*)
\]

(24)

From KKT condition (i), we get

\[
f(x) + \left[ \sum_{j=1}^{s} \mu_j^+ \nabla h_j(x^*) + \sum_{j=1}^{m} \zeta^+_i \nabla g_i(x^*) \right] \eta(x, x^*) = f(x) - \nabla f(x^*) \eta(x, x^*)
\]

(25)

Since \( f \) is also invex at \( x^* \) on \( X \) with respect to the same function \( \eta \) and by invex definition with the feasibility of the problem \( P \) at \( x^* \), we have

\[
f(x) - \nabla f(x^*) \eta(x, x^*) \geq f(x^*)
\]

(26)

\[
f(x^*) + c \left[ \sum_{j=1}^{s} |h_j(x^*)| + \sum_{i=1}^{m} g_i^+(x^*) \right] = f(x^*) + c \left( \sum_{j=1}^{s} \ln \left( (h_j(x^*))^2 + 1 \right) \right) + \sum_{i=1}^{m} \ln \left( (g_i^+(x^*))^2 + 1 \right)
\]

(27)

\[
P(x^*, c) = f(x^*) + c \left( \sum_{j=1}^{s} \ln \left( (h_j(x^*))^2 + 1 \right) \right) + \sum_{i=1}^{m} \ln \left( (g_i^+(x^*))^2 + 1 \right)
\]

(28)

From Equation (20) to Equation (28), we conclude that the following inequality holds for any \( x \in X \)

\[
P(x, c) \geq P(x^*, c)
\]

This shows that \( x^* \) is a minimizer of the penalized optimization problem (19). This completes the proof of the theorem.

**Corollary 4.** (This is a corollary to Theorem 3) Let \( x^* \) be a point in the optimization problem \( P \), and the suitable constraint qualifications are satisfied at \( x^* \) for the problem \( P \). Moreover, suppose that the function \( f, g_i, i \in I(x^*) \), \( h_j, j \in J^+ \) are invex at \( x^* \) on \( X \) with respect to the same function \( \eta \) and the function \( h_j, j \in J^- \) are invave at \( x^* \) on \( X \) with respect to the same function \( \eta \). If \( c \) is sufficiently large (setting \( c \geq \max \{ \zeta_i^+, i \in I, |\mu_j^+|, j \in J \} \)), where both \( \mu_j^+, j = 1, 2, \ldots, s, \zeta_i^+, i = 1, 2, \ldots, m \) are Lagrange multipliers associated with the equality and inequality constraints \( h_j \) and \( g \), respectively, then \( x^* \) is also a minimizer of the penalized problem (19).

**Proposition 5.** Let the point \( x^* \) be a minimizer of the logarithmic penalty function (15) in the penalized problem (19). Then the following inequality holds for all \( x \in F \).

\[
f(x) \geq f(x^*)
\]

**Proof.** Suppose that \( x^* \) minimizes the logarithmic penalty function (15) in the penalized problem (19). Since \( x^* \) is an optimal solution in the problem (19), then the following inequality holds for all \( x \in F \).

\[
f(x) + c \left( \sum_{j=1}^{s} \ln \left( (h_j(x))^2 + 1 \right) + \sum_{i=1}^{m} \ln \left( (g_i^+(x))^2 + 1 \right) \right) \geq f(x^*) + c \left( \sum_{j=1}^{s} \ln \left( (h_j(x^*))^2 + 1 \right) + \sum_{i=1}^{m} \ln \left( (g_i^+(x^*))^2 + 1 \right) \right)
\]

Therefore, for all \( x \in X \)

\[
f(x) + c \left( \sum_{j=1}^{s} \ln \left( (h_j(x))^2 + 1 \right) + \sum_{i=1}^{m} \ln \left( (g_i^+(x))^2 + 1 \right) \right) \geq f(x^*)
\]

Hence, by definition of \( g_i^+(x) \) the following inequality is true for all \( x \in F \).

\[
f(x) \geq f(x^*)
\]

This concludes the proof of the proposition.

Now, the following theorem establishes contradictory results without any assumption that the functions constituting the problem \( P \) are invex.

**Theorem 6.** Let \( x^* \) be a minimizer of logarithmic penalty function (15) in the penalized problem (19). Furthermore, let the inequality \( P(x, c) \geq P(x^*, c) \) be satisfied for any \( c \geq c^* \) and for all \( x \in F \). Then \( x^* \) is also an optimal solution in the original optimization problem \( P \).

**Proof.**

Let us assume that \( x^* \) minimizes the logarithmic penalty function (15). By Proposition 5, the following inequality holds for all \( x \in F \). i.e.

\[
f(x) \geq f(x^*)
\]

(29)

The inequality (29) indicates that the function’s values are bounded below on the feasible set \( F \) in the problem \( P \).

Now, our goal is to prove that \( x^* \) is also optimal in the original optimization problem \( P \). Accordingly, we will first show that \( x^* \) is feasible in the original problem \( P \).
By contradiction, suppose that $x^*$ is not feasible to the problem (P). Hence, by (1), we have

$$
\sum_{j=1}^{s} \ln \left( \frac{1^i}{c^2} \right) + \sum_{j=1}^{m} \ln \left( \frac{1^i}{c^2} \right) > 0
$$

Based on the hypothesis, $P(x, c) \geq P(x^*, c)$ is satisfied for any $c \geq c^*$ and for all $x \in F$.

Let us assume that $x$ be any feasible solution in the optimization problem (P). Hence, by (1), we have

$$
f(x) + \frac{f(x)}{c^2} > f(x^*) + \frac{f(x^*)}{c^2}
$$

We set

$$
\frac{f(x) - f(x^*)}{c^2} = \max \left\{ \frac{\ln \left( \frac{1^i}{c^2} \right) + \sum_{j=1}^{m} \ln \left( \frac{1^i}{c^2} \right)}{c^2}, \frac{\ln \left( \frac{1^i}{c^2} \right) + \sum_{j=1}^{m} \ln \left( \frac{1^i}{c^2} \right)}{c^2} \right\}
$$

Therefore,

$$
f(x^*) + \frac{f(x^*)}{c^2} > f(x) + \frac{f(x)}{c^2}
$$

But

$$
f(x) = f(x) + \frac{f(x)}{c^2}
$$

Combining (33) and (34) we get

$$
f(x^*) + \frac{f(x^*)}{c^2} > f(x) + \frac{f(x)}{c^2} + \frac{f(x)}{c^2}
$$

which is a contradiction to (30). Hence, $x^*$ is feasible in the problem (P) and its optimality follows from (29). Obviously, Theorem 6 has been well-established even though no invexity assumption is required on the functions constituting the optimization problem (P). Now, let us consider the case when invexity assumption is imposed on the problem constituted by invex functions, and we are to show that there exists the edge of a penalty parameter $c$, above which the result is true.

**Theorem 7.** Let $x^*$ be a minimizer of the logarithmic penalty function (1) in the penalized optimization problem (19). Moreover, let the functions $f, g_i, i \in I$, $h_j, j \in J^*(\bar{x}) := \{ j \in J : \pi_j > 0 \}$ are invex at any optimal point $\bar{x}$ in the problem (P) on $X$ with respect to the same function $\eta$, and the function $h_j, j \in J^*(\bar{x}) := \{ j \in J : \pi_j < 0 \}$ are invex at any optimal point $\bar{x}$ in (P) on $X$ with respect to the same function $\eta$. The set of all feasible points in the original optimization problem is compact, and the penalty parameter $c$ is considered to be sufficiently large (we set $c > \max \{ \zeta_i, i \in I \}$). Therefore, $x^*$ is also optimal to the problem (P).

**Proof.** We are required to show that $x^*$ is also optimal in the optimization problem (P). First, we need to show that $x^*$ is among the feasible points in the problem (P). By contradiction, let us assume that $x^*$ is not feasible in the problem (P). By weierstrass theorem, $\tilde{f}$ admits its minimum $\tilde{x}$ on a compact set $\tilde{F}$ (since $\tilde{f}$ is a continuous function bounded below on the set $\tilde{F}$). Therefore, the original optimization problem (P) has an optimal solution $\tilde{x}$. Based on this hypothesis, the functions $f, g_i, i \in I(\tilde{x})$, $h_j, j \in J^*(\tilde{x})$ are invex at $\tilde{x}$ on $X$ with respect to the same function $\eta$ and the function $h_j, j \in J^*(\tilde{x})$ are incave at $\tilde{x}$ on $X$ with respect to the same function $\eta$. Therefore, by definitions of invex and incave functions, we have the following:

$$
f(x^*) - f(\tilde{x}) \geq \nabla f(\tilde{x}) \cdot \eta(x^*, \tilde{x})
$$

$$
g_i(x^*) - g_i(\tilde{x}) \geq \nabla g_i(\tilde{x}) \cdot \eta(x^*, \tilde{x}), \quad i \in I(\tilde{x})
$$

$$
h_j(x^*) - h_j(\tilde{x}) \geq \nabla h_j(\tilde{x}) \cdot \eta(x^*, \tilde{x}), \quad j \in J^*(\tilde{x})
$$

$$
h_j(x^*) - h_j(\tilde{x}) \leq \nabla h_j(\tilde{x}) \cdot \eta(x^*, \tilde{x}), \quad j \in J^*(\tilde{x})
$$

Since $\tilde{x}$ is optimal in the optimization problem (P), there exist Lagrange multipliers $\zeta_i \in \mathbb{R}_{+}$ and $\pi_i \in \mathbb{R}$ such that the KKT necessary optimality conditions (i)- (iii) hold. Then, by (19)–(21),

$$
\zeta_i g_i(x^*) - \zeta_i g_i(\tilde{x}) \geq \zeta_i \nabla g_i(\tilde{x}) \cdot \eta(x^*, \tilde{x}), \quad i \in I
$$

$$
\pi_j h_j(x^*) - \pi_j h_j(\tilde{x}) \geq \pi_j \nabla h_j(\tilde{x}) \cdot \eta(x^*, \tilde{x}), \quad j \in J
$$

Upon adding the both sides of (35), (39) and (40), we get

$$
f(x^*) - f(\tilde{x}) + \sum_{j=1}^{m} \pi_j h_j(x^*) - \sum_{j=1}^{m} \pi_j h_j(\tilde{x})
$$

$$
\geq \left[ \nabla f(\tilde{x}) + \sum_{j=1}^{m} \pi_j \nabla h_j(\tilde{x}) + \sum_{i=1}^{m} \zeta_i \nabla g_i(\tilde{x}) \right] \cdot \eta(x^*, \tilde{x})
$$
By KKT necessary optimality conditions (i), we obtain

\[ f(x^*) + \sum_{j=1}^{s} \mu_j^* |h_j(x^*)| + \sum_{i=1}^{m} \xi_i g_i(x^*) \geq f(x^*) \]

\[ + \sum_{j=1}^{s} \nabla h_j(x^*) + \sum_{i=1}^{m} \nabla g_i(x^*) \geq f(x) \]  \hspace{1cm} (41)

\[ + \sum_{j=1}^{s} \nabla h_j(x) + \sum_{i=1}^{m} \nabla g_i(x) \]

By the feasibility of \( x \) and KKT condition (ii), we have

\[ \sum_{j=1}^{s} f(\bar{x}) + \sum_{j=1}^{s} \nabla h_j(\bar{x}) + \sum_{i=1}^{m} \xi_i g_i(\bar{x}) \]

\[ = f(\bar{x}) + c \left( \sum_{j=1}^{s} \ln \left( (h_j(\bar{x}))^2 + 1 \right) \right) \]

\[ + \sum_{i=1}^{m} \ln \left( (g_i(\bar{x}))^2 + 1 \right) \]

Since \( \sum_{j=1}^{s} \nabla h_j(\bar{x}) = 0 \), then

\[ \sum_{j=1}^{s} \ln \left( (h_j(\bar{x}))^2 + 1 \right) = 0. \]  \hspace{1cm} (43)

Moreover,

\[ \sum_{i=1}^{m} \xi_i g_i(x) = 0. \]  \hspace{1cm} (44)

Hence, by (41)–(44),

\[ f(x^*) + \sum_{j=1}^{s} \mu_j^* |h_j(x^*)| + \sum_{i=1}^{m} \xi_i g_i(x^*) \geq f(\bar{x}) \]  \hspace{1cm} (45)

Since by assumption, the penalty parameter \( c \) is assumed to be sufficiently large (we set \( c > \max \left\{ \xi_i, i \in L, \ n_j, \ j \in J \right\} \)), also \( x^* \) is assumed to be not feasible by hypothesis in the original optimization problem (P), we expect one of \( g_i(x^*) \) and \( h_j(x^*) \) to be non-zero, and the following inequality holds:

\[ f(x^*) + c \left( \sum_{j=1}^{s} \ln \left( (h_j(x^*))^2 + 1 \right) + \sum_{i=1}^{m} \ln \left( (g_i(x^*))^2 + 1 \right) \right) \]

\[ \geq f(x^*) + \sum_{j=1}^{s} \mu_j^* |h_j(x^*)| + \sum_{i=1}^{m} \xi_i g_i(x^*) > f(x) \]

Then, the feasibility of \( x^* \), the subsequent expression becomes

\[ f(x^*) + c \left( \sum_{j=1}^{s} \ln \left( (h_j(x^*))^2 + 1 \right) + \sum_{i=1}^{m} \ln \left( (g_i(x^*))^2 + 1 \right) \right) \]

\[ > f(\bar{x}) + c \left( \sum_{j=1}^{s} \ln \left( (h_j(\bar{x}))^2 + 1 \right) + \sum_{i=1}^{m} \ln \left( (g_i(\bar{x}))^2 + 1 \right) \right) . \]

Then, Equation (19) that defined the penalized optimization problem based on the logarithmic penalty function, it follows that the following strict inequality holds:

\[ P(x^*, c) > P(\bar{x}, c) \]

which is a contradiction, since \( x^* \) is a minimizer of the penalized problem (19). Therefore, we have proved that \( x^* \) is feasible in the original optimization problem (P). Hence, from Proposition 5, the optimality of \( x^* \) in the problem (P) follows. Thus, this has established the conclusion of the theorem.

The following result follows directly from Corollary 4 and Theorem 7.

**Corollary 8.** Let the assumptions of Corollary 4 and Theorem 7 be satisfied. Then, the sets of optimal solutions of the original optimization problem (P) and the minimizers of its associated logarithmic penalized problem (19) coincide.

The theoretical result has shown that there exists a finite value of the penalty parameter \( c \) under relevant invexity assumptions imposed on the functions constituting the problem (P) and some suitable constrained qualification, that will retrieve an optimal solution of the problem (P) via the penalized logarithmic minimization problem (19).

Now, we can exhibit the result via the following invex optimization problem.

**Example 1.** (Antczak, 2009). Consider the following optimization problem:

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

\[ \text{s.t. } g_1(x) = -\ln(x_1 + 1) \leq 0, \]

\[ g_2(x) = -(x_2^2 + 1) \arctan x_2 \leq 0, \]

\[ X = \left\{ (x_1, x_2) \in \mathbb{R}^2 : x_1 > -1 \right\} \]

It can be verified that the objective function \( f \) and the constraints \( g_1 \) and \( g_2 \) are invex on the set \( X \) with respect to the function \( \eta \) defined by

\[ \eta(x, u) = \left( \frac{g_1(x)}{g_1(u)} \right), \]

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

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

via Definition 2. Clearly, \( F = \{ (x_1, x_2) \in X : x_1 \geq 0 \land x_2 \geq 0 \} \).

The considered optimization problem can be solved by constructing a single unconstrained problem via a logarithmic penalty function method defined in (1). Therefore, the following unconstrained problem can be obtained:

\[ P(x, c) = \ln(x_1 + 1) + (x_2^2 + 1) \arctan x_2 \]

\[ + \ln \left( \max (0, -\ln(x_1 + 1)) \right)^2 + 1 \]

\[ + \ln \left( \max (0, -(x_2^2 + 1) \arctan x_2) \right)^2 + 1 \]
Observe that $x^* = (0, 0)$ is feasible in the considered penalized problem and KKT necessary optimality conditions (i)-(iii) are fulfilled at $x^* = (0, 0)$ with multipliers $\bar{\lambda}_1 = \bar{\lambda}_2 = 1$. We set the penalty parameter $c > \max\{\bar{\lambda}_i, i \in I\}$, as stated in Theorem 3. Figures 1–3 represented the considered problem in its original form, absolute value penalty function form and logarithmic penalty function form, respectively. The results of all the three forms indicate that $x^* = (0, 0)$ is an optimal solution in the considered non-linear programming problem. Furthermore, the stated sufficient conditions for convex optimization problem in Theorem 9.3.1 (Bazaraa et al., 1991) are not applicable, since not all functions constituting the problem (P) are convex but the functions constituting the problem (P) are invex with respect to the same function $\eta$ on the set $X$. Therefore, it follows that sufficient conditions in Theorem 3 are essential for the non-convex optimization problem.

**Example 2.** Consider the following optimization problem:

$$\min f(x) = x^3$$
$$g(x) = -x - 1 \leq 0.$$

Obviously, the objective function is not an invex function with respect to any function $\eta$ on $R$ defined by $\eta : R \times R \rightarrow R$. Note that $F = \{x \in R : x \geq -1\}$.
and \( x^* = -1 \) is an optimal solution in the considered problem. Therefore, we use the proposed logarithmic penalty function (15). Then we have the following single unconstrained optimization problem:

\[
\min P(x, c) = x^3 + \log(\max(0, -x - 1))^2 + 1
\]

\( P(x, c) \) does not have a minimizer at \( x^* = -1 \) for any values of \( c > 0 \) as depicted in Figures 4–6, where \( c = 10 \). Therefore, no equivalence exists between the original optimization problem, its exact penalized optimization problem via the absolute value penalty function and its penalized problem with proposed logarithmic penalty function since the downward growth of \( f \) exceeds that of the \( g \) upward growth at \( x^* \). Indeed, the invexity assumption is essential to ensure the equivalence between the sets of optimal solutions in the original optimization problem and its transformed proposed unconstrained problem with logarithmic penalty function.

5. Numerical test

In this section, some small problems were tested with different classifications to justify the proposed logarithmic penalty function efficiency. All the problems were selected from Hock and Schittkowski (1981) collections set of continuous problems with equality and inequality constraints. The following
Table 1. Results obtained by solving test problems with mixed (equality and inequality) constraints from Hock-Schittkowski collection.

| P-NAME | n | m | AVP | GLP | Iteration | Objective value | Classification |
|--------|---|---|-----|-----|-----------|----------------|---------------|
| HS014  | 2 | 2 | 08  | 11  | 2         | 1.3935E + 00  | 1.6227E + 00  | QQR           |
| HS032  | 3 | 5 | 18  | 09  | 1         | 1.0000E + 00  | 2.2761E-04   | QPR           |
| HS041  | 4 | 5 | 04  | 20  | 2         | 1.9259E + 00  | 1.9486E + 00  | QPR           |
| HS071  | 4 | 6 | 06  | 44  | 3         | 1.7014E + 01  | 1.8791E + 01  | QPR           |
| HS073  | 4 | 7 | 03  | 27  | 4         | 2.9894E + 01  | 3.1564E + 01  | LGR           |
| HS074  | 4 | 9 | 05  | 55  | 5         | 5.1265E + 03  | 6.7530E + 03  | PGR           |
| HS075  | 4 | 9 | 05  | 54  | 6         | 5.1744E + 03  | 6.7537E + 03  | PGR           |
| HS080  | 5 | 8 | **  | 16  | 7         | 5.3950E-02    | Unbounded     | GPR           |
| HS081  | 5 | 8 | **  | 36  | 8         | 5.3950E-02    | Unbounded     | GPR           |

Figure 5. Constrained problem via absolute value penalty function.

Figure 6. Constrained problem via logarithmic penalty function.

The table comes up with comparisons of the objective values for the constrained problems of some selected from Hock-Schittkowski, absolute value penalty function method and the proposed logarithmic penalty function method. The experiment was implemented via quasi-Newton algorithm using the
The classification of each problem provides more detailed information regarding the problem and consists of letters, i.e. OCD.

- O-information about objective function
- C-information about constraint function
- D-information about the regularity of the problem

6. Results and discussions

Based on the results obtained in the first part, it was observed that the problem solved by Antczak (2009) reiterated that invexity assumptions are essential to establish an equivalence between the sets of optimal solutions in the original optimization problem and its associated penalized problem as depicted in the Figures 1–3 for the original problem, penalized problem via absolute value penalty function and penalized problem via proposed logarithmic penalty function, respectively; \( x^* = (0, 0) \) happens to be the optimal solution for all the three methods. On the other hand, in the subsequent problem the objective is not an invex function on \( R \) with respect to any function \( \eta \), and this yields to different optimal solutions as illustrated in Figures 4–6 for the original constrained problem, penalized problem via exact absolute value penalty function and penalized problem by means of proposed logarithmic penalty function, respectively.

However, in the second part, it can be observed that for the problems with classifications QQR, QPR, PLR and PPR, the proposed logarithmic penalty proved to be more efficient compared to the original formulation of the constrained optimization problem and penalized problem with exact absolute value penalty function. Nevertheless, for QPR \( \text{fminunc} \) stopped because it cannot decrease the objective function along the current search direction, even though its number of iterations is higher than that of GLP, while in the remaining GLP it admits more numbers of iterations compared to exact absolute value penalty function formulation; the same conclusions go for LGI and PGR.

Moreover, in the last two problems with classification GPR, the absolute value penalty function method via the same algorithm, problems HS080 and HS081 turn out to be unbounded while the proposed logarithmic penalty function method was able to solve the problems.

7. Conclusion

This paper reviewed some of the existing penalty function methods for solving constrained optimization problem and proposed a continuously differentiable penalty function called logarithmic penalty function method; the proposed approach is the hybridization of logarithmic penalty function and modified Courant-Beltrami penalty function for equality and inequality constraints, respectively. The equivalence between the sets of the optimal solutions for the original optimization problem constituted by invex functions with respect to the same function \( \eta \) and its associated penalized problem via the proposed penalty function has been established. According to the results obtained, the concept of invexity is essential to determine the equivalence between the sets of the optimal solution in the original problem and its associated penalized problem.

Few problems with mixed (equality and inequality) constraints from the Hock-Schittkowski collection were solved by means of quasi-Newton algorithm via routine function \( \text{fminunc} \) (MATLAB2018a) as in Table 1; the objective values obtained for the proposed logarithmic penalty function turn out to be more advantageous to decision makers from a practical point of view, even though there is a need to address some issues regarding the number of iterations (Table 2). Future work will focus on inventing a suitable algorithm for this type of formulation and also to solve practical problems.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Mansur Hassan http://orcid.org/0000-0001-7736-2059
Adam Baharum http://orcid.org/0000-0002-1541-982X

References

Antczak, T. (2009). Exact penalty functions method for mathematical programming problems involving invex
