



Duality in convex optimization for the hyperbolic augmented Lagrangian

Lennin Mallma Ramirez¹  · Nelson Maculan² · Adilson Elias Xavier¹ · Vinicius Layter Xavier³

Received: 24 October 2023 / Accepted: 2 March 2024 / Published online: 11 April 2024
© The Author(s) 2024

Abstract

We guarantee the strong duality and the existence of a saddle point of the hyperbolic augmented Lagrangian function (HALF) in convex optimization. To guarantee these results, we assume a set of convexity hypothesis and the Slater condition. Finally, we computationally illustrate our theoretical results obtained in this work.

Keywords Hyperbolic augmented Lagrangian · Nonlinear programming · Constrained convex optimization · Saddle point · Duality

Mathematics Subject Classification 90C30 · 90C25 · 90C46 · 65K05

Nelson Maculan, Adilson Elias Xavier, and Vinicius Layter Xavier have contributed equally to this work.

✉ Lennin Mallma Ramirez
lenninmr@gmail.com

Nelson Maculan
maculan@cos.ufrj.br

Adilson Elias Xavier
adilson.xavier@gmail.com

Vinicius Layter Xavier
viniciuslx@ime.uerj.br

- ¹ Systems Engineering and Computer Science Program (COPPE), Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
- ² Systems Engineering and Computer Science Program-Applied Mathematics (COPPE/IM), Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
- ³ Department of Statistics, Institute of Mathematics and Statistics, Graduate Program in Computational Sciences, Rio de Janeiro State University, Rio de Janeiro, Brazil

1 Introduction

Throughout this work, we are interested in studying the following optimization problem:

$$(P) \quad x^* \in X^* = \operatorname{argmin}\{f(x) \mid x \in S\},$$

where

$$S = \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, \quad i = 1, \dots, m\}$$

is the convex feasible set of the problem (P), the function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex, $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, m$, is concave and we also assume that f and g_i are continuously differentiable.

The augmented Lagrangian algorithms solve convex problem (P), see [11] and [12]. Recently, it was guaranteed that the hyperbolic augmented Lagrangian algorithm (HALA) also solves problem (P), see [9]. In [9], it is guaranteed that a subsequence generated by HALA converges toward a Karush–Kuhn–Tucker (KKT) point. The HALA minimizes the HALF (this minimization problem is known as the subproblem generated by the augmented Lagrangian algorithm). With the solution found in the subproblem, the Lagrange multipliers will be updated through an update formula. The theory of duality is a current topic and this theory is studied in more general spaces, see [20]. In this work, we are interested in developing the duality theory for HALF in the Euclidean space.

The main result of our work is: we guarantee the strong duality for the HALF for the convex case. In this way, we assure a solution to the primal and dual problems. With these results, we can also note that HALF has similar properties to Log-Sigmoid Lagrangian function (LSLF), see [12], modified Frisch function (MFF), and modified Carroll function (MCF), these last two functions are studied in [11].

The work is organized as follows: in Sect. 2, we present some basic results, the hyperbolic penalty function, and some of its properties; in Sect. 3, we present the HALF; in Sect. 4, we develop the duality theory for HALF, and in Sect. 5, we present a set of computational illustrations to verify our theoretical results.

Notation

Throughout this work, \mathbb{R} is the Euclidean space, $\mathbb{R}_+ = \{a \in \mathbb{R} : a \geq 0\}$, $\mathbb{R}_{++} = \{a \in \mathbb{R} : a > 0\}$ and $\|\cdot\|$ is the Euclidean norm.

2 Preliminaries

Henceforth, we consider the following assumption:

A. Slater constraint qualification holds, i.e., there exists $\hat{x} \in S$ which satisfies $g_i(\hat{x}) > 0$, $i = 1, \dots, m$.

The Lagrangian function associated with the problem (P) is $L : \mathbb{R}^n \times \mathbb{R}_+^m \rightarrow \mathbb{R}$, defined as

$$L(x, \lambda) = f(x) - \sum_{i=1}^m \lambda_i g_i(x),$$

where $\lambda_i \geq 0, i = 1, \dots, m$, are called Lagrange multipliers. The dual function $\Phi : \mathbb{R}_+^m \rightarrow \mathbb{R}$, is defined as

$$\Phi(\lambda) = \inf_{x \in \mathbb{R}^n} L(x, \lambda), \tag{2.1}$$

and the dual problem consists of finding (D):

$$\lambda^* \in \Lambda^* = \operatorname{argmax}\{\Phi(\lambda) \mid \lambda \in \mathbb{R}_+^m\}.$$

2.1 Hyperbolic penalty

The hyperbolic penalty method, introduced in [16], is meant to solve the problem (P). The penalty method adopts the hyperbolic penalty function (HPF) as follows:

$$P(y, \lambda, \tau) = -\lambda y + \sqrt{(\lambda y)^2 + \tau^2}, \tag{2.2}$$

where $P : \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_{++} \rightarrow \mathbb{R}$. The graphic representation of $P(y, \lambda, \tau)$ is as shown in Fig. 1. Now, in the present work, we consider $\tau > 0$ fixed.

Remark 2.1 The HPF is originally proposed in [16] and also studied in [17]. In these works, the following properties are important for HPF:

- (a) $P(y, \lambda, \tau)$ is asymptotically tangent to the straight lines $r_1(y) = -2\lambda y$ and $r_2(y) = 0$ for $\tau > 0$.
- (b) $P(y, \lambda, 0) = 0$, for $y \geq 0$ and $P(y, \lambda, 0) = -2\lambda y$, for $y < 0$.

Due to the properties (a) and (b), the HPF perform a smoothing of the penalty studied by Zangwill, see [19].

For more details of the HPF, see [16] and [17]. In particular, we use the following properties of the function P:

- (P0) $P(y, \lambda, \tau)$ is k -times continuously differentiable for any positive integer k for $\tau > 0$.
- (P1) $P(0, \lambda, \tau) = \tau$, for $\tau > 0$ and $\lambda \geq 0$.
- (P2) $P(y, \lambda, \tau)$ is strictly a decreasing function of y , i.e.,

$$\nabla_y P(y, \lambda, \tau) = -\lambda \left(1 - \frac{\lambda y}{\sqrt{(\lambda y)^2 + \tau^2}} \right) < 0, \text{ for } \tau > 0, \lambda > 0.$$

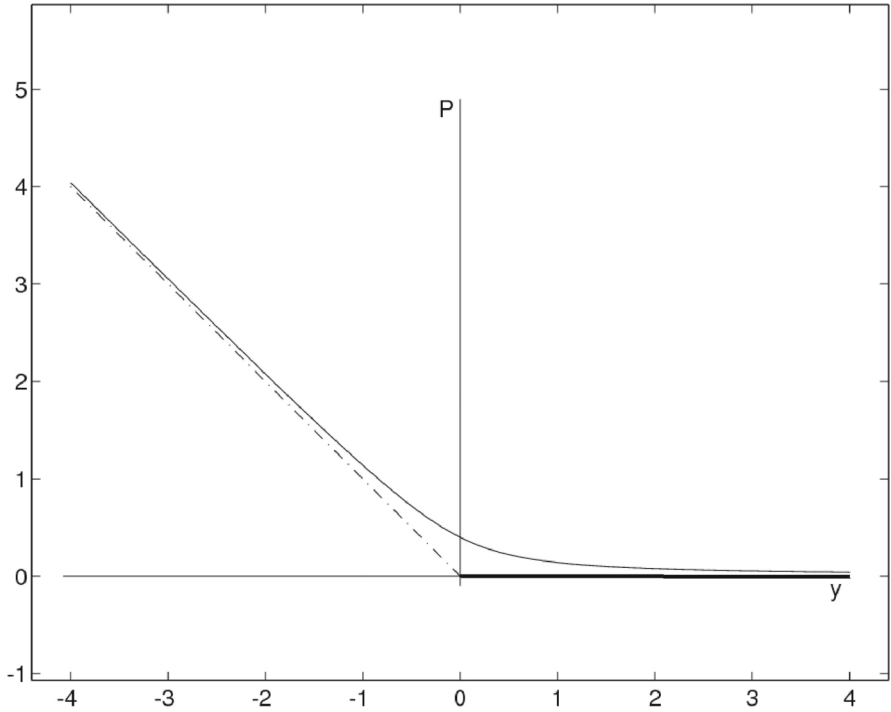


Fig. 1 Hyperbolic penalty function

Remark 2.2 For any $\lambda \geq 0$, $y \geq 0$ and $\tau > 0$. We have

$$-\lambda y - \sqrt{(\lambda y)^2 + \tau^2} < 0 < -\lambda y + \sqrt{(\lambda y)^2 + \tau^2}. \tag{2.3}$$

Considering the definition of the function P in (2.3), we have $P(y, \lambda, \tau) > 0$.

3 Hyperbolic augmented Lagrangian function

We define the HALF of problem (P) by $L_H : \mathbb{R}^n \times \mathbb{R}_{++}^m \times \mathbb{R}_{++} \rightarrow \mathbb{R}$,

$$L_H(x, \lambda, \tau) = f(x) + \sum_{i=1}^m P(g_i(x), \lambda_i, \tau),$$

which can be written as

$$L_H(x, \lambda, \tau) = f(x) + \sum_{i=1}^m \left(-\lambda_i g_i(x) + \sqrt{(\lambda_i g_i(x))^2 + \tau^2} \right), \tag{3.1}$$

where τ is the penalty parameter and a fixed value. Note that this function belongs to class C^∞ if the involved functions $f(x)$ and $g_i(x)$, $i = 1, \dots, m$, are too.

Proposition 3.1 *Let us assume that if $f(x)$ and all $g_i(x) \in C^2$, and that $f(x)$ is strictly convex, then $L_H(x, \lambda, \tau)$ is strictly convex in \mathbb{R}^n for any fixed $\lambda > 0$ and $\tau > 0$.*

Proof We only need to prove that the Hessiana of L_H is defined positive. Let $\lambda = (\lambda_1, \dots, \lambda_m) > 0$ and $\tau > 0$ be fixed. The Hessian of $L_H(x, \lambda, \tau)$ is

$$\begin{aligned} \nabla_{xx}^2 L_H(x, \lambda, \tau) &= \nabla_{xx}^2 f(x) - \sum_{i=1}^m \lambda_i \nabla_{xx}^2 g_i(x) \\ &\quad + \sum_{i=1}^m \left(\frac{(\lambda_i)^2}{\sqrt{(\lambda_i g_i(x))^2 + \tau^2}} - \frac{(\lambda_i)^4 g_i^2(x)}{((\lambda_i g_i(x))^2 + \tau^2)^{\frac{3}{2}}} \right) \nabla_x g_i(x) \nabla_x^T g_i(x) \\ &\quad + \sum_{i=1}^m \frac{(\lambda_i)^2 g_i(x)}{\sqrt{(\lambda_i g_i(x))^2 + \tau^2}} \nabla_{xx}^2 g_i(x). \end{aligned} \tag{3.2}$$

In (3.2), the $\nabla_{xx}^2 g_i(x)$ is factored, so we can rewrite (3.2) as follows:

$$\begin{aligned} \nabla_{xx}^2 L_H(x, \lambda, \tau) &= \nabla_{xx}^2 f(x) - \sum_{i=1}^m \lambda_i \left(1 - \frac{\lambda_i g_i(x)}{\sqrt{(\lambda_i g_i(x))^2 + \tau^2}} \right) \nabla_{xx}^2 g_i(x) \\ &\quad + \sum_{i=1}^m \left(\frac{(\lambda_i)^2}{\sqrt{(\lambda_i g_i(x))^2 + \tau^2}} - \frac{(\lambda_i)^4 g_i^2(x)}{((\lambda_i g_i(x))^2 + \tau^2)^{\frac{3}{2}}} \right) \nabla_x g_i(x) \nabla_x^T g_i(x). \end{aligned} \tag{3.3}$$

On the other hand, we know that we have

$$\left((\lambda_i g_i(x))^2 + \tau^2 \right) \lambda_i^2 > \lambda_i^2 (\lambda_i g_i(x))^2.$$

The above inequality can be written as

$$\frac{\lambda_i^2}{((\lambda_i g_i(x))^2 + \tau^2)^{\frac{1}{2}}} - \frac{\lambda_i^4 g_i^2(x)}{((\lambda_i g_i(x))^2 + \tau^2)^{\frac{3}{2}}} > 0. \tag{3.4}$$

Replacing (3.4) in (3.3) and since $-\lambda_i \left(1 - \frac{\lambda_i g_i(x)}{\sqrt{(\lambda_i g_i(x))^2 + \tau^2}} \right) < 0$, in (3.3), we get that, $\nabla_{xx}^2 L_H(x, \lambda, \tau) > 0$, for $\lambda > 0$ and $\tau > 0$ fixed. □

Remark 3.1 Henceforth, we are going to assume that there is a solution for the subproblem generated by HALA. Now, considering this assumption of existence of a solution and considering the Proposition 3.1 for any $\lambda > 0$ and $\tau > 0$, then there exists a unique minimizer for problem (P). This minimizer is defined as

$$\check{x} = \check{x}(\lambda, \tau) = \operatorname{argmin} \{ L_H(x, \lambda, \tau) \mid x \in \mathbb{R}^n \}.$$

4 Duality

In this section, we adapt the classic results already existing in the literature (Chapter 9 of [10] and Section 7 of [11]) for our HALF. The following result is also verified by MFF and MCF, see [11].

Proposition 4.1 *Consider the convex problem (P). Assume the assumption A holds. Then, $x^* \in S$ is a solution of problem (P) for any $\tau > 0$ if and only if:*

(i) *There exists a vector $\lambda^* \geq 0$ such that*

$$\lambda_i^* g_i(x^*) = 0, \quad i = 1, \dots, m \quad \text{and} \quad L_H(x, \lambda^*, \tau) \geq L_H(x^*, \lambda^*, \tau), \quad \forall x \in \mathbb{R}^n. \tag{4.1}$$

(ii) *The pair (x^*, λ^*) is a saddle point of L_H , that is,*

$$L_H(x, \lambda^*, \tau) \geq L_H(x^*, \lambda^*, \tau) \geq L_H(x^*, \lambda, \tau), \quad \forall x \in \mathbb{R}^n, \quad \forall \lambda \in \mathbb{R}_+^m. \tag{4.2}$$

Proof (\Rightarrow) Let any $\tau > 0$ be fixed. Assume x^* is a solution for convex problem (P) satisfying the assumption A. Then the system

$$\begin{aligned} f(x) - f(x^*) &< 0, \\ -g_i(x) &< 0, \quad i = 1, \dots, m, \end{aligned}$$

has no solution in \mathbb{R}^n . Hence, by the Proper Separation Theorem (see, Theorem 2.26 (iv) of Dhara and Dutta [3]), there exists a vector $(\tilde{\lambda}, \hat{\lambda}) \neq (0, 0) \in \mathbb{R} \times \mathbb{R}^m$ such that

$$\tilde{\lambda} (f(x) - f(x^*)) - \sum_{i=1}^m \hat{\lambda}_i g_i(x) \geq 0,$$

for all $x \in \mathbb{R}^n$. We can rewrite the inequality above as

$$\tilde{\lambda} (f(x) - f(x^*)) \geq \sum_{i=1}^m \hat{\lambda}_i g_i(x), \tag{4.3}$$

for all $x \in \mathbb{R}^n$. Now, we follow an analysis similar to Theorem 4.2 of [3], so by A, we have that there exists $\lambda_i^* = \frac{\hat{\lambda}_i}{\tilde{\lambda}}, \quad i = 1, \dots, m,$ with $\tilde{\lambda} > 0$. Then, by (4.3) we have

$$f(x) - f(x^*) \geq \sum_{i=1}^m \lambda_i^* g_i(x), \tag{4.4}$$

for all $x \in \mathbb{R}^n$. In particular, (4.4) holds for $x = x^*$. So, we have

$$0 \geq \sum_{i=1}^m \lambda_i^* g_i(x^*). \tag{4.5}$$

On the other hand, since $g_i(x^*) \geq 0$ and $\lambda_i^* \geq 0$ for $i = 1, \dots, m$, then by (4.5), we obtain

$$\lambda_i^* g_i(x^*) = 0, \quad i = 1, \dots, m, \tag{4.6}$$

which holds, thus, we have the first part of (4.1).

Now, we are interested in proving the second part of (4.1). From (4.6) and (4.4), we have

$$f(x^*) - \sum_{i=1}^m \lambda_i^* g_i(x^*) = f(x^*) \leq f(x) - \sum_{i=1}^m \lambda_i^* g_i(x), \tag{4.7}$$

for all $x \in \mathbb{R}^n$. Now, since we have (4.6), we can obtain

$$\sum_{i=1}^m \sqrt{(\lambda_i^* g_i(x^*))^2 + \tau^2} \leq \sum_{i=1}^m \sqrt{(\lambda_i^* g_i(x))^2 + \tau^2}. \tag{4.8}$$

Considering (4.7) and (4.8), we have

$$L_H(x, \lambda^*, \tau) \geq L_H(x^*, \lambda^*, \tau), \quad \forall x \in \mathbb{R}^n, \tag{4.9}$$

which concludes the proof of (4.1).

Now, we are interested in verifying item (ii), first we will prove that $L_H(x^*, \lambda^*, \tau) = f(x^*) + m\tau$. Indeed, we consider (4.6) in the definition of function L_H (see (3.1)), so we have

$$L_H(x^*, \lambda^*, \tau) = f(x^*) + m\tau. \tag{4.10}$$

On the other hand, x^* is feasible, i.e.,

$$g_i(x^*) \geq 0, \quad i = 1, \dots, m. \tag{4.11}$$

By applying the property P2 of the HPF in (4.11), we obtain

$$P(g_i(x^*), \lambda_i, \tau) \leq P(0, \lambda_i, \tau), \quad i = 1, \dots, m. \tag{4.12}$$

By applying property P1 on the right side of expression (4.12) and by performing the sum of 1 to m in the inequality, it follows

$$\sum_{i=1}^m P(g_i(x^*), \lambda_i, \tau) \leq \sum_{i=1}^m \tau = m\tau.$$

Adding $f(x^*)$ to both sides of the expression, we obtain

$$f(x^*) + \sum_{i=1}^m P(g_i(x^*), \lambda_i, \tau) \leq f(x^*) + m\tau. \tag{4.13}$$

By definition of L_H , (4.13) becomes

$$L_H(x^*, \lambda, \tau) \leq f(x^*) + m\tau. \tag{4.14}$$

Now, by (4.14) and (4.10), we have

$$L_H(x^*, \lambda, \tau) \leq f(x^*) + m\tau = L_H(x^*, \lambda^*, \tau). \tag{4.15}$$

Finally, from (4.9) and (4.15), there exists $\lambda^* \geq 0$ such that the primal-dual solution (x^*, λ^*) is a saddle point of L_H , $\forall x \in \mathbb{R}^n$ and $\tau > 0$.

(\Leftarrow) We assume that (x^*, λ^*) is a saddle point of L_H , so (4.2) holds. Then, for all $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}_+^m$ and for any $\tau > 0$ fixed, we have

$$\begin{aligned} f(x^*) - \sum_{i=1}^m \lambda_i g_i(x^*) + \sum_{i=1}^m \sqrt{(\lambda_i g_i(x^*))^2 + \tau^2} &= L_H(x^*, \lambda, \tau) \\ &\leq L_H(x^*, \lambda^*, \tau) = f(x^*) - \sum_{i=1}^m \lambda_i^* g_i(x^*) + \sum_{i=1}^m \sqrt{(\lambda_i^* g_i(x^*))^2 + \tau^2}. \end{aligned} \tag{4.16}$$

This relation (4.16) is possible only if $g_i(x^*) \geq 0$. If this relation is violated (*i.e.*, $g_i(x^*) < 0$) for some index i , we can choose λ_i sufficiently large so that (4.16) becomes false. So, x^* is feasible for problem (P).

We will prove the complementarity condition of (4.1). So again, by (4.16), since that $\lambda_i \geq 0$, $i = 1, \dots, m$, and in particular taking $\lambda_i = 0$, $i = 1, \dots, m$, in (4.16), we obtain

$$\sum_{i=1}^m \left((\lambda_i^* g_i(x^*))^2 + \tau^2 + 2\tau \lambda_i^* g_i(x^*) \right) \leq \sum_{i=1}^m \left((\lambda_i^* g_i(x^*))^2 + \tau^2 \right).$$

Thus, $\sum_{i=1}^m \lambda_i^* g_i(x^*) \leq 0$, and since $\lambda_i^* \geq 0$ and $g_i(x^*) \geq 0$, $i = 1, \dots, m$, it must be true that

$$\lambda_i^* g_i(x^*) = 0, \quad i = 1, \dots, m. \tag{4.17}$$

By (4.17) and the definition of L_H , we obtain

$$L_H(x^*, \lambda^*, \tau) = f(x^*) + m\tau. \tag{4.18}$$

From the definition of saddle point, we know that $L_H(x, \lambda^*, \tau) \geq L_H(x^*, \lambda^*, \tau)$. By (4.18) and by the definition of L_H , we can write

$$\begin{aligned} f(x^*) + m\tau &= L_H(x^*, \lambda^*, \tau) \leq L_H(x, \lambda^*, \tau) \\ &= f(x) + \sum_{i=1}^m P(g_i(x), \lambda_i^*, \tau). \end{aligned} \tag{4.19}$$

On the other hand, once again considering the property P2 of HPF, for any feasible point x , i.e., $g_i(x) \geq 0, i = 1, \dots, m$, we will carry out a similar approach as was done on equation (4.11)-(4.13), thus, obtaining

$$f(x) + \sum_{i=1}^m P(g_i(x), \lambda_i^*, \tau) \leq f(x) + m\tau. \tag{4.20}$$

Now, we replace (4.20) in (4.19), so it follows $f(x^*) \leq f(x)$ whenever x is feasible. Therefore, x^* is a global optimal solution of (P). □

Let us consider the following definitions. Let

$$F_\tau(x) = \sup_{\lambda \geq 0} L_H(x, \lambda, \tau).$$

Then, $F_\tau(x) = f(x) + m\tau$ if $g_i(x) \geq 0, i = 1, \dots, m$ and $F_\tau(x) = \infty$, otherwise. Therefore, we can consider the following problem:

$$x^* = \operatorname{argmin} \{ F_\tau(x) \mid x \in \mathbb{R}^n \}, \tag{4.21}$$

that is the problem (P) reduces to solving (4.21).

Let

$$\phi_\tau(\lambda) = \inf_{x \in \mathbb{R}^n} L_H(x, \lambda, \tau)$$

(possibly $\phi_\tau(\lambda) = -\infty$ for some λ) and consider the following dual problem of (P) that consists of finding

$$\lambda^* = \operatorname{argmax} \{ \phi_\tau(\lambda) \mid \lambda \geq 0 \}. \tag{4.22}$$

In the following result, we are going to verify the weak duality.

Proposition 4.2 *Let x be a feasible solution to problem (P) and let λ be a feasible solution to problem (4.22). Then,*

$$\phi_\tau(\lambda) \leq F_\tau(x) = f(x) + m\tau, \quad \forall x \in S, \forall \lambda \in \mathbb{R}_+^m.$$

Proof For any feasible x and λ , we have the weak duality. Indeed, by the definition of ϕ_τ , we have

$$\begin{aligned} \phi_\tau(\lambda) &= \inf_{w \in \mathbb{R}^n} L_H(w, \lambda, \tau) = \inf_{w \in \mathbb{R}^n} \left\{ f(w) + \sum_{i=1}^m P(g_i(w), \lambda_i, \tau) \right\} \\ &\leq \inf_{w \in S} \left\{ f(w) + \sum_{i=1}^m P(g_i(w), \lambda_i, \tau) \right\} \\ &= f(x) + \sum_{i=1}^m P(g_i(x), \lambda_i, \tau). \end{aligned} \tag{4.23}$$

Since we know that x is feasible, we immediately have $g_i(x) \geq 0, i = 1, \dots, m$. Then, from the property P2 of the HPF, we have the following expressions

$$\sum_{i=1}^m P(g_i(x), \lambda_i, \tau) \leq \sum_{i=1}^m P(0, \lambda_i, \tau).$$

Now, applying property P1 on the right side of the previous inequality, and adding $f(x)$, to both sides of the inequality above, we have

$$f(x) + \sum_{i=1}^m P(g_i(x), \lambda_i, \tau) \leq f(x) + m\tau. \tag{4.24}$$

Replacing (4.24) in (4.23), we finally obtain

$$\phi_\tau(\lambda) \leq f(x) + m\tau, \quad \forall x \in S, \forall \lambda \in \mathbb{R}_+^m. \tag{4.25}$$

□

If \hat{x} and $\hat{\lambda}$ are feasible solutions of the primal and dual problems and $F_\tau(\hat{x}) = \phi_\tau(\hat{\lambda})$, then $\hat{x} = x^*$ and $\hat{\lambda} = \lambda^*$. From Remark 3.1, with the smoothness of $f(x)$ and $g_i(x), i = 1, \dots, m$, we ensure the smoothness for the dual function $\phi_\tau(\lambda)$.

Theorem 4.1 *The problem (P) is considered. The assumption **A** is verified. Then, the existence of a solution of problem (P) implies that the problem (4.22) has a solution and*

$$\phi_\tau(\lambda^*) = f(x^*) + m\tau, \quad \text{for any } \tau > 0. \tag{4.26}$$

Proof Let x^* be a solution of problem (P). From **A**, we get $\lambda^* \geq 0$, so that (4.1) is verified. So, we have

$$\begin{aligned} \phi_\tau(\lambda^*) &= \min_{x \in \mathbb{R}^n} L_H(x, \lambda^*, \tau) = L_H(x^*, \lambda^*, \tau) \\ &\geq L_H(x^*, \lambda, \tau) \geq \min_{x \in \mathbb{R}^n} L_H(x, \lambda, \tau) = \phi_\tau(\lambda), \quad \forall \lambda \geq 0. \end{aligned}$$

Therefore, $\phi_\tau(\lambda^*) = \max \{ \phi_\tau(\lambda) \mid \lambda \in \mathbb{R}_+^m \}$. In this way, $\lambda^* \in \mathbb{R}_+^m$ is a solution of the dual problem and since we have $L_H(x^*, \lambda^*, \tau) = f(x^*) + m\tau$, (4.26) holds. \square

Proposition 4.3 *Suppose that (4.26) holds, for the viable points x^* and λ^* . Then, x^* is a solution of the problem (P) and λ^* is a solution of the dual problem (4.22).*

Proof Let $g_i(x^*) \geq 0, i = 1, \dots, m$, with $x^* \in S, \lambda_i^* \geq 0, i = 1, \dots, m$ and (4.26) with $\tau > 0$ be fixed. Then from (4.25), where x and λ are viable, we can obtain the following

$$f(x) + m\tau \geq \phi_\tau(\lambda^*) = f(x^*) + m\tau \geq \phi_\tau(\lambda),$$

i.e., x^* is solution of the problem (P) and λ^* is solution of (4.22), which corresponds to the validity of the strong duality. \square

5 Computational illustration

We are going to use HALA (see [9]) to guarantee the theory proposed in this work. The program was compiled by the GNU Fortran compiler version 4:7.4.0-1ubuntu2.3. The numerical experiments were conducted on a notebook with the operating system Ubuntu 18.04.5, CPU i7-3632QM and 8GB RAM. The unconstrained minimization tasks were carried out by means of a Quasi-Newton algorithm employing the BFGS updating formula, with the function VA13 from HSL library [6]. The algorithm stopped when the solution was viable (feasible) and the absolute value of the difference of the two consecutive solutions $\|x^k - x^{k-1}\|$ was less than 10^{-5}

We are going to take advantage of this section to make some comparisons of our algorithm HALA (see Table 18) with respect to the following algorithms:

- Alg1 = [4], a truncated Newton method;
- Alg2 = [5], a primal–dual interior-point method;
- Alg3 = [14], an interior-point algorithm;
- Alg4 = [13], a QP-free method;
- Alg5 = [2], a primal–dual feasible interior-point method;
- Alg6 = [18], a feasible sequential linear equation algorithm;
- Alg7 = [15], an inexact first-order method
- Alg8 = [7], a feasible direction interior-point technique;
- Alg9 = [1], an interior-point algorithm.

For completeness reasons, we are going to present HALA:

Algorithm 1 Hyperbolic Augmented Lagrangian Algorithm

Step 1. Let $k := 0$ (initialization). Take initial values $\lambda^0 = (\lambda_1^0, \dots, \lambda_m^0) \in \mathbb{R}_{++}^m, \tau \in \mathbb{R}_{++}$.

Step 2. Solve the unconstrained minimization problem (primal update):

$$\begin{aligned} x^{k+1} &\in \operatorname{argmin}_{x \in \mathbb{R}^n} L_H(x, \lambda^k, \tau) \\ &= \operatorname{argmin}_{x \in \mathbb{R}^n} \left\{ f(x) + \sum_{i=1}^m \left(-\lambda_i^k g_i(x) + \sqrt{(\lambda_i^k g_i(x))^2 + \tau^2} \right) \right\}. \end{aligned}$$

Step 3. Update the Lagrange multipliers (dual update):

$$\lambda_i^{k+1} = \lambda_i^k \left(1 - \frac{\lambda_i^k g_i(x^{k+1})}{\sqrt{(\lambda_i^k g_i(x^{k+1}))^2 + \tau^2}} \right), \quad i = 1, \dots, m. \tag{5.1}$$

Step 4. If the pair (x^{k+1}, λ^{k+1}) satisfies the stopping criteria, stop.

Step 5. $k := k + 1$. Go to Step 2.

With the following examples proposed in the book [8], we are going to verify the strong duality. In each example, the value of m means the total number of restrictions. Also, in all the examples starting points are considered so that assumption **A** is verified.

Example 5.1 Problem 1 (HS1).

$$\begin{aligned} \min_{x \in \mathbb{R}^2} f(x) &= 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \\ \text{s.t. } g_1(x) &= x_2 + 1.5 \geq 0. \end{aligned}$$

Starting with $x^0 = (-2, 1)$ (feasible), $f(x^0) = 909$ and $m = 1$. The minimum value is $f(x^*) = 0$ at the optimal solution $x^* = (1, 1)$.

Example 5.2 Problem 30 (HS30).

$$\begin{aligned} \min_{x \in \mathbb{R}^3} f(x) &= x_1^2 + x_2^2 + x_3^2 \\ \text{s.t. } g_1(x) &= x_1^2 + x_2^2 - 1 \geq 0, \\ g_2(x) &= x_1 - 1 \geq 0, \\ g_3(x) &= 10 - x_1 \geq 0, \\ g_4(x) &= x_2 + 10 \geq 0, \\ g_5(x) &= 10 - x_2 \geq 0, \\ g_6(x) &= x_3 + 10 \geq 0, \\ g_7(x) &= 10 - x_3 \geq 0. \end{aligned}$$

Starting with $x^0 = (1, 1, 1)$ (feasible), $f(x^0) = 3$ and $m = 7$. The minimum value is $f(x^*) = 1$ at the optimal solution $x^* = (1, 0, 0)$.

Example 5.3 Problem 66 (HS66).

$$\begin{aligned} \min_{x \in \mathbb{R}^3} f(x) &= 0.2x_3 - 0.8x_1 \\ \text{s.t. } g_1(x) &= x_2 - e^{x_1} \geq 0, \\ g_2(x) &= x_3 - e^{x_2} \geq 0, \\ g_3(x) &= x_1 \geq 0, \\ g_4(x) &= x_2 \geq 0, \\ g_5(x) &= x_3 \geq 0, \\ g_6(x) &= 100 - x_1 \geq 0, \\ g_7(x) &= 100 - x_2 \geq 0, \\ g_8(x) &= 10 - x_3 \geq 0. \end{aligned}$$

Starting with $x^0 = (0, 1.05, 2.9)$ (feasible), $f(x^0) = 0.58$ and $m = 8$. The minimum value is $f(x^*) = 0.5181632741$ at the optimal solution $x^* = (0.1841264879, 1.202167873, 3.327322322)$.

Example 5.4 Problem 76 (HS76).

$$\begin{aligned} \min_{x \in \mathbb{R}^4} f(x) &= x_1^2 + 0.5x_2^2 + x_3^2 + 0.5x_4^2 - x_1x_3 + x_3x_4 - x_1 - 3x_2 + x_3 - x_4 \\ \text{s.t. } g_1(x) &= 5 - x_1 - 2x_2 - x_3 - x_4 \geq 0, \\ g_2(x) &= 4 - 3x_1 - x_2 - 2x_3 + x_4 \geq 0, \\ g_3(x) &= x_2 + 4x_3 - 1.5 \geq 0, \\ g_4(x) &= x_1 \geq 0, \\ g_5(x) &= x_2 \geq 0, \\ g_6(x) &= x_3 \geq 0, \\ g_7(x) &= x_4 \geq 0. \end{aligned}$$

Starting with $x^0 = (0.5, 0.5, 0.5, 0.5)$ (feasible), $f(x^0) = -1.25$ and $m = 7$. The minimum value is $f(x^*) = -4.681818181$ at the optimal solution $x^* = (0.2727273, 2.090909, -0.26E-10, 0.5454545)$.

Example 5.5 Problem 100 (HS100).

$$\begin{aligned} \min_{x \in \mathbb{R}^7} f(x) &= (x_1 - 10)^2 + 5(x_2 - 12)^2 + x_3^4 + 3(x_4 - 11)^2 + 10x_5^6 \\ &\quad + 7x_6^2 + x_7^4 - 4x_6x_7 - 10x_6 - 8x_7 \\ \text{s.t. } g_1(x) &= 127 - 2x_1^2 - 3x_2^4 - x_3 - 4x_4^2 - 5x_5 \geq 0, \\ g_2(x) &= 282 - 7x_1 - 3x_2 - 10x_3^2 - x_4 + x_5 \geq 0, \\ g_3(x) &= 196 - 23x_1 - x_2^2 - 6x_6^2 + 8x_7 \geq 0, \\ g_4(x) &= -4x_1^2 - x_2^2 + 3x_1x_2 - 2x_3^2 - 5x_6 + 11x_7 \geq 0. \end{aligned}$$

Table 1 Example 5.1

k	x_1	x_2	$f(x)$	$L_H(x, \lambda, \tau)$	Via
0	- 0.200000000E+01	0.100000000E+01	0.909000000E+03	0.909000000E+03	0
1	0.100000000E+01	0.100000000E+01	0.976202768E-23	0.200017780E-11	0
2	0.100000000E+01	0.100000000E+01	0.976202768E-23	0.999999800E-05	0

Table 2 Example 5.1

k	$g_1(x)$	
	Via	λ_1
0	0	0.100000000E+02
1	0	0.800470801E-12
2	0	0.800470641E-12

Starting with $x^0 = (1, 2, 0, 4, 0, 1, 1)$ (feasible), $f(x^0) = 714$ and $m = 4$. The minimum value is $f(x^*) = 680.6300573$ at the optimal solution $x^* = (2.330499, 1.951372, -0.4775414, 4.365726, -0.6244870, 1.038131, 1.594227)$.

Example 5.6 Problem 11 (HS11).

$$\begin{aligned} \min_{x \in \mathbb{R}^2} f(x) &= (x_1 - 5)^2 + x_2^2 - 25 \\ \text{s.t. } g_1(x) &= -x_1^2 + x_2 \geq 0, \end{aligned}$$

Starting with $x^0 = (4.9, 0.1)$ (not feasible), $f(x^0) = -24.98$ and $m = 1$. The minimum value is $f(x^*) = -8.498464223$.

5.1 Results

For each table, the letter N indicates the name of the problem, λ is the multiplier Lagrange, x is the primal variable, $f(x)$ is the value of the objective function, $g_i(x)$ are the constraints of each problem, $L_H(\cdot, \cdot, \cdot)$ is the value of the HALF and via = viable = feasible where, in each iteration, the obtained point can be viable, then its value is “0 = yes” or the point can be inviable, then the value is “1 = not” and τ is the penalty parameter. In all the examples, we will use $\tau = 0.10E-04$. Let us analyze the examples.

- Example 5.1: The HALA solves this example even though function f is non-convex, see Tables 1 and 2. The time used is 0.000250 s.
- Example 5.2: the function f is strictly convex. From Table 3, we can see that in iteration 2, the Theorem 4.1 can be verified, i.e., we have the following:

$$f(x^*) + m\tau = 1.00000000 + (7)(0.00001) = 1.00007$$

Table 3 Example 5.2

k	x_1	x_2	x_3	$f(x)$	$L_H(x, \lambda, \tau)$	Via
0	0.100000000E+01	0.100000000E+01	0.100000000E+01	0.300000000E+01	0.300001000E+01	0
1	0.100000179E+01	- 0.773363648E- 10	- 0.477975865E- 10	0.100000359E+01	0.100000755E+01	0
2	0.100000000E+01	- 0.198472639E- 09	0.162237191E- 09	0.100000000E+01	0.100007000E+01	0

Table 4 Example 5.2

k	$\underline{g_1(x)}$		$\underline{g_2(x)}$		$\underline{g_3(x)}$	
	Via	λ_1	Via	λ_2	Via	λ_3
0	0	0.100000000E+02	0	0.100000000E+02	0	0.100000000E+02
1	0	0.367214575E+00	0	0.126557312E+01	0	0.610622664E-13
2	0	0.367214283E+00	0	0.126557139E+01	0	0.610622630E-13

Table 5 Continuation of Table 4

k	$\underline{g_4(x)}$		$\underline{g_5(x)}$		$\underline{g_6(x)}$		$\underline{g_7(x)}$	
	via	λ_4	via	λ_5	via	λ_6	via	λ_7
0	0	0.100000000E+02	0	0.100000000E+02	0	0.100000000E+02	0	0.100000000E+02
1	0	0.499600361E-13	0	0.499600361E-13	0	0.499600361E-13	0	0.499600361E-13
2	0	0.499600336E-13	0	0.499600336E-13	0	0.499600336E-13	0	0.499600336E-13

and

$$\phi_\tau(\lambda^*) = L_H(x^*, \lambda^*, \tau) = 1.00007,$$

then, $\phi_\tau(\lambda^*) = f(x^*) + m\tau$. So, $x^* = (0.100000000E+01, 0.100000000E+01)$ is the solution of the primal problem and from Tables 4 and 5, we can see the λ^* is the solution of the dual problem in the iteration 2. The time used is 0.000434 s.

- Example 5.3: the function f is linear. From Table 6, we can see that in iteration 3 the Theorem 4.1 can be verified, i.e., we have the following:

$$f(x^*) + m\tau = 0.518163274 + (8)(0.00001) = 0.518243274$$

and

$$\phi_\tau(\lambda^*) = L_H(x^*, \lambda^*, \tau) = 0.518243274,$$

then, $\phi_\tau(\lambda^*) = f(x^*) + m\tau$. So, x^* is the solution of the primal problem and from Tables 7 and 8, we can see the λ^* is the solution of the dual problem in the iteration 3. The time used is 0.000804 s.

- Example 5.4: the function f is strictly convex. From Table 9, we can see that in iteration 2, the Theorem 4.1 can be verified, i.e., we have the following:

$$f(x^*) + m\tau = -4.68181818 + (7)(0.00001) = -4.68174818$$

and

$$\phi_\tau(\lambda^*) = L_H(x^*, \lambda^*, \tau) = -4.68174818,$$

Table 6 Example 5.3

k	x_1	x_2	x_3	$f(x)$	$L_H(x, \lambda, \tau)$	Via
0	0.000000000E+00	0.105000000E+01	0.290000000E+01	0.580000000E+00	0.580010000E+00	0
1	0.184125306E+00	0.120216905E+01	0.332733118E+01	0.518165991E+00	0.518168851E+00	0
2	0.184126486E+00	0.120216787E+01	0.332732231E+01	0.518163274E+00	0.518243274E+00	0
3	0.184126486E+00	0.120216787E+01	0.332732231E+01	0.518163274E+00	0.518243274E+00	0

Table 7 Example 5.3

k	$g_1(x)$		$g_2(x)$		$g_3(x)$		$g_4(x)$	
	Via	λ_1	Via	λ_2	Via	λ_3	Via	λ_4
0	0	0.100000000E+02	0	0.100000000E+02	0	0.100000000E+02	0	0.100000000E+02
1	0	0.665464690E+00	0	0.200000037E+00	0	0.147483137E-09	0	0.345945494E-11
2	0	0.665464311E+00	0	0.199999981E+00	0	0.147482736E-09	0	0.345945351E-11
3	0	0.665463933E+00	0	0.199999924E+00	0	0.147482336E-09	0	0.345945207E-11

Table 8 Continuation of Table 7

k	$g_5(x)$		$g_6(x)$		$g_7(x)$		$g_8(x)$	
	Via	λ_5	Via	λ_6	via	λ_7	Via	λ_8
0	0	0.100000000E+02	0	0.100000000E+02	0	0.100000000E+02	0	0.100000000E+02
1	0	0.450750548E-12	0	0.111022302E-14	0	0.000000000E+00	0	0.113242749E-12
2	0	0.450750480E-12	0	0.111022301E-14	0	0.000000000E+00	0	0.113242740E-12
3	0	0.450750413E-12	0	0.111022300E-14	0	0.000000000E+00	0	0.113242731E-12

then, $\phi_\tau(\lambda^*) = f(x^*) + m\tau$. So, x^* is the solution of the primal problem and from Tables 10 and 11, we can see the λ^* is the solution of the dual problem in the iteration 2. The time used is 0.000652 s.

- Example 5.5: the function f is convex. From Table 12, we can see that in iteration 2, the Theorem 4.1 can be verified, i.e., we have the following:

$$f(x^*) + m\tau = 680.630057 + (4)(0.00001) = 680.630097$$

and

$$\phi_\tau(\lambda^*) = L_H(x^*, \lambda^*, \tau) = 680.630097,$$

then, $\phi_\tau(\lambda^*) = f(x^*) + m\tau$. The optimal value x^* is reported in the Table 12 and the optimal value λ^* is reported in the Table 13. The time used is 0.002140 s.

In Table 18, we can see that HALA is more efficient in the sense that it uses fewer iterations with respect to the other algorithms. We can observe in the computational results that the HALA remains in the viable region in all the examples. On the other hand, despite being the theory developed in this work on convexity hypothesis, our algorithm shows in the Example 5.1 that it can also solve non-convex problems.

A computational experiment with update of parameter “ τ ”

We show a computational experiment, considering a condition to update parameter τ . So, we present our Algorithm 2. In this algorithm we will additionally assume that the sequence $\{\tau^k\}$ is bounded.

- Example 5.6.

This example is solved with Algorithm 1 and Algorithm 2.

Table 9 Example 5.4

k	x_1	x_2	x_3	x_4	$f(x)$	$L_H(x, \lambda, \tau)$	Via
0	0.500000000E+00	0.500000000E+00	0.500000000E+00	0.500000000E+00	-0.125000000E+01	-0.125000000E+01	0
1	0.272727650E+00	0.209090766E+01	0.147253122E-05	0.545452356E+00	-0.468181418E+01	-0.468180958E+01	0
2	0.272727273E+00	0.209090909E+01	0.413220517E-10	0.545454545E+00	-0.468181818E+01	-0.468174818E+01	0

Table 10 Example 5.4

k	$g_1(x)$		$g_2(x)$		$g_3(x)$		$g_4(x)$	
	Via	λ_1	Via	λ_2	Via	λ_3	Via	λ_4
0	0	0.100000000E+02	0	0.100000000E+02	0	0.100000000E+02	0	0.100000000E+02
1	0	0.454545522E+00	0	0.186739513E-11	0	0.143196566E-10	0	0.672217837E-10
2	0	0.454545455E+00	0	0.186739456E-11	0	0.143196445E-10	0	0.672216605E-10

Table 11 Continuation of Table 10

k	$g_5(x)$		$g_6(x)$		$g_7(x)$	
	Via	λ_5	Via	λ_6	Via	λ_7
0	0	0.100000000E+02	0	0.100000000E+02	0	0.100000000E+02
1	0	0.114352972E-11	0	0.172728506E+01	0	0.168043357E-10
2	0	0.114352972E-11	0	0.172728506E+01	0	0.168043357E-10

Algorithm 2 Hyperbolic Augmented Lagrangian Algorithm

Step 1. Let $k := 0$ (initialization). Take initial values $\lambda^0 = (\lambda_1^0, \dots, \lambda_m^0) \in \mathbb{R}_{++}^m$, $0 < \alpha < 1$ and $\tau^0 \in \mathbb{R}_{++}$.

Step 2. Solve the unconstrained minimization problem (primal update):

$$\begin{aligned}
 &x^{k+1} \in \operatorname{argmin}_{x \in \mathbb{R}^n} L_H(x, \lambda^k, \tau^k) \\
 &= \operatorname{argmin}_{x \in \mathbb{R}^n} \left\{ f(x) + \sum_{i=1}^m \left(-\lambda_i^k g_i(x) + \sqrt{(\lambda_i^k g_i(x))^2 + (\tau^k)^2} \right) \right\}.
 \end{aligned}$$

Step 3. Update the Lagrange multipliers (dual update):

$$\lambda_i^{k+1} = \lambda_i^k \left(1 - \frac{\lambda_i^k g_i(x^{k+1})}{\sqrt{(\lambda_i^k g_i(x^{k+1}))^2 + (\tau^k)^2}} \right), \quad i = 1, \dots, m. \tag{5.2}$$

Step 4. Update of parameter τ . If x^{k+1} is feasible, define

$$\tau^{k+1} = \tau^k.$$

Else, define

$$\tau^{k+1} = \alpha \tau^k.$$

Step 5. $k := k + 1$. Go to Step 2.

For both algorithms, we are considering the same initial values: $x^0 = (4.9, 0.1)$, $\lambda^0 = 10$, $\tau^0 = 0.01$ and $\alpha = 0.2$ (for the case of Algorithm 2). The algorithm stopped when the solution was viable (feasible) and the absolute value of the difference of the two consecutive solutions $\|x^k - x^{k-1}\|$ was less than 10^{-3} . Thus, we have the following observations:

Table 12 Example 5.5

k	x_1	x_2	x_3	x_4	x_5
0	0.100000000E+01	0.200000000E+01	0.000000000E+00	0.400000000E+01	0.000000000E+00
1	0.233049916E+01	0.195137237E+01	-0.477541322E+00	0.436572623E+01	-0.624486971E+00
2	0.233049937E+01	0.195137237E+01	-0.477541393E+00	0.436572623E+01	-0.624486971E+00
x_6	x_7	$f(x)$	$L_H(x, \lambda, \tau)$		
0.100000000E+01	0.100000000E+01	0.714000000E+03	0.714000000E+03	0	Via
0.103813098E+01	0.159422676E+01	0.680630061E+03	0.680630065E+03	0	
0.103813102E+01	0.159422671E+01	0.680630057E+03	0.680630097E+03	0	

Table 13 Example 5.5

k	$g_1(x)$		$g_2(x)$		$g_3(x)$		$g_4(x)$	
	Via	λ_1	Via	λ_2	Via	λ_3	Via	λ_4
0	0	0.100000000E+02	0	0.100000000E+02	0	0.100000000E+02	0	0.100000000E+02
1	0	0.113971988E+01	0	0.000000000E+00	0	0.000000000E+00	0	0.368614695E+00
2	0	0.113971989E+01	0	0.000000000E+00	0	0.000000000E+00	0	0.368614517E+00

Table 14 Example 5.6 with $\tau^0 = 0.01$

k	x_1	x_2	$f(x)$	$L_H(x, \lambda, \tau)$	via
1	0.123478414E+01	0.152463713	- 0.849863113E+01	- 0.848846557E+01	1
2	0.123477887E+01	0.152467882	- 0.849846436E+01	- 0.848846422E+01	1
3	0.123477887E+01	0.152467882	- 0.849846436E+01	- 0.848846422E+01	1
4	0.123477887E+01	0.152467882	- 0.849846436E+01	- 0.848846422E+01	1
5	0.123477887E+01	0.152467882	- 0.849846436E+01	- 0.848846422E+01	1
6	0.123477887E+01	0.152467882	- 0.849846436E+01	- 0.848846422E+01	1
7	0.123477887E+01	0.152467882	- 0.849846436E+01	- 0.848846422E+01	1
8	0.123477887E+01	0.152467882	- 0.849846436E+01	- 0.848846422E+01	1
9	0.123477887E+01	0.152467882	- 0.849846436E+01	- 0.848846422E+01	1
10	0.123477887E+01	0.152467882	- 0.849846436E+01	- 0.848846422E+01	1
11	0.123477887E+01	0.152467882	- 0.849846436E+01	- 0.848846422E+01	1
12	0.123477872E+01	0.152467887	- 0.849846303E+01	- 0.848846422E+01	0

Table 15 Example 5.6 with $\tau^0 = 0.01$

k	$g_1(x)$	
	Via	λ_1
1	1	0.304925717E+01
2	1	0.304930063E+01
3	1	0.304934410E+01
4	1	0.304938757E+01
5	1	0.304943104E+01
6	1	0.304947452E+01
7	1	0.304951799E+01
8	1	0.304956146E+01
9	1	0.304960494E+01
10	1	0.304964842E+01
11	1	0.304969189E+01
12	0	0.304932925E+01

Table 16 Example 5.6 with $\tau^0 = 0.01$

k	x_1	x_2	$f(x)$	$L_H(x, \lambda, \tau)$	Via
1	0.123478414E+01	0.152463713E+01	- 0.849863113E+01	- 0.84884655E+01	1
2	0.123474541E+01	0.152459622E+01	- 0.849846421E+01	- 0.84964642E+01	0

Table 17 Example 5.6 with $\tau^0 = 0.01$

k	Via	$g_1(x)$	
		λ_1	τ
1	1	0.304925717E+01	0.002
2	0	0.304925379E+01	0.002

Table 18 Iterations

N	HALA	Alg1	Alg2	Alg3	Alg4	Alg5	Alg6	Alg7	Alg8	Alg9
HS1	2	18	34	32	40	24	36	260	36	27
HS30	2	3	8	11	7	7	10	7	11	10
HS66	3	12	12	13		11	11	23	5	20
HS76	2	28	9	11	10		9	23	7	12
HS100	2	18	10	11	15	9	14	99	13	14

(a) Algorithm 1 (with τ fixed).

The results can be seen on the Tables 14 and 15. We can observe that this algorithm uses 12 iterations to solve the problem.

(b) Algorithm 2 (with a rule to update τ).

The results can be seen on the Tables 16 and 17. We can observe that this algorithm uses 2 iterations to solve the problem.

After seeing (a) and (b), we can decide that Algorithm 2 solves Example 5.6, using fewer iterations compared to Algorithm 1. This motivates us to investigate new rules to update parameter τ .

Acknowledgements Lennin Mallma Ramirez was supported by PDR10/FAPERJ E-26/205.684/2022; Nelson Maculan by COPPETEC Foundation, Brazil and the Vinicius Layer Xavier is thankful to Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) for their financial support E-26/2021—Auxílio Básico à Pesquisa, APQ1. We the authors want to thank Igor Pereira dos Santos Pereira (PEQ/COPPE/UFRJ, Brazil), for helping with the computational experiments.

Declarations

Conflict of interest The authors declare that there are no competing interest.

References

1. Akrotirianakis, I., Rustem, B.: Globally convergent interior-point algorithm for nonlinear programming. *J. Optim. Theory Appl.* **125**(3), 497–521 (2005)
2. Bakhtiari, S., Tits, A.L.: A simple primal-dual feasible interior-point method for nonlinear programming with monotone descent. *Comput. Optim. Appl.* **25**, 17–38 (2003)
3. Dhara, A., Dutta, J.: *Optimality Conditions in Convex Optimization: A finite-Dimensional View*. CRC Press, Boca Raton (2012)
4. Di Pillo, G., Lucidi, S., Palagi, L.: A truncated Newton method for constrained optimization. In: Di Pillo, G., Giannesi, F. (eds.) *Nonlinear Optimization and Related Topics*. Kluwer Academic, Dordrecht (2000)
5. Griva, I., Shanno, D.F., Vanderbei, R.J., Benson, H.Y.: Global convergence of a primal-dual interior point method for nonlinear programming. *Algorithmic Oper. Res.* **3**, 12–29 (2008)
6. Harwell Subroutine Library: A collection of Fortran codes for large scale scientific computation. <https://www.hsl.rl.ac.uk/> (2002)
7. Herskovits, J.: Feasible direction interior-point technique for nonlinear optimization. *J. Optim. Theory Appl.* **99**, 121–146 (1998)
8. Hock, W., Schittkowski, K.: *Test Examples for Nonlinear Programming Codes*. Lecture Notes in Economics and Mathematical Systems, vol. 187. Springer, Berlin (1981)
9. Mallma-Ramirez, L., Maculan, N., Xavier, A.E.: Convergence analysis of the hyperbolic augmented Lagrangian algorithm. Technical Report, PESC-COPPE, Federal University of Rio de Janeiro. ES-2991/21. <https://www.cos.ufrj.br/index.php/pt-BR/publicacoes-pesquisa/details/15/2991> (2021)
10. Polyak, B.T.: *Introduction to Optimization*. Optimization Software Inc, New York (1987)
11. Polyak, R.A.: Modified barrier functions: theory and methods. *Math. Program.* **54**, 177–222 (1992)
12. Polyak, R.A.: Log-sigmoid multipliers method in constrained optimization. *Ann. Oper. Res.* **101**, 427–460 (2001)
13. Qi, H.D., Qi, L.: A new QP-free, globally convergent, locally superlinearly convergent algorithm for inequality constrained optimization. *SIAM J. Optim.* **11**, 113–132 (2000)
14. Vanderbei, R.J., Shanno, D.F.: An interior-point algorithm for nonconvex nonlinear programming. *Comput. Optim. Appl.* **13**, 231–252 (1999)
15. Wang, H., Zhang, F., Wang, J., Rong, Y.: An inexact first-order method for constrained nonlinear optimization. *Optim. Methods Softw.* (2018). <https://doi.org/10.1080/10556788.2020.1712601>
16. Xavier, A.E.: *Penalização Hiperbólica- Um Novo Método para Resolução de Problemas de Otimização* (master's thesis). Federal University of Rio de Janeiro \ COPPE, Rio de Janeiro (1982)
17. Xavier, A.E.: Hyperbolic penalty: a new method for nonlinear programming with inequalities. *Int. Trans. Oper. Res.* **8**, 659–671 (2001)
18. Yang, Y.F., Li, D.H., Qi, L.: A feasible sequential linear equation method for inequality constrained optimization. *SIAM J. Optim.* **13**(4), 1222–1244 (2003)
19. Zangwill, W.I.: Non-linear programming via penalty function. *Manag. Sci.* **13**(5), 344–358 (1967)
20. Zhao, L., Zhu, D.L.: On iteration complexity of a first-order primal-dual method for nonlinear convex cone programming. *J. Oper. Res. Soc. China* (2021). <https://doi.org/10.1007/s40305-021-00344-x>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.