



# Article An Efficient Penalty Method without a Line Search for Nonlinear Optimization

Assma Leulmi

Department of Mathematics, Ferhat Abbas University of Setif-1, Setif 19137, Algeria; as\_smaleulmi@yahoo.fr or assma.leulmi@univ-setif.dz

**Abstract:** In this work, we integrate some new approximate functions using the logarithmic penalty method to solve nonlinear optimization problems. Firstly, we determine the direction by Newton's method. Then, we establish an efficient algorithm to compute the displacement step according to the direction. Finally, we illustrate the superior performance of our new approximate function with respect to the line search one through a numerical experiment on numerous collections of test problems.

**Keywords:** interior point methods; logarithmic penalty method; line search; approximate functions; nonlinear optimization

MSC: 90C25; 90C30; 90C51

#### 1. Introduction

The nonlinear optimization is a fundamental subject in the modern optimization literature. It focuses on the problem of optimizing an objective function in the presence of inequality and/or equality constraints. Furthermore, the optimization problem is obviously linear if all the functions are linear, otherwise it is called a nonlinear optimization problem.

This research field is motivated by the fact that it arises in various problems encountered in practice, such as business administration, economics, agriculture, mathematics, engineering, and physical sciences.

In our knowledge, Frank and Wolfe are the deans in nonlinear optimization problems. They established a powerful algorithm in [1] to solve them. Later, they used another method in [2] based on the application of the Simplex method on the nonlinear problem after converting it to a linear one.

This pioneer work inspired many authors to propose and develop several methods and techniques to solve this class of problems. We refer to [3,4] for interior point methods to find the solution of nonlinear optimization problems with a high dimension.

In order to make this theory applicable in practice, other methods are designed on the linear optimization history, among robust algorithms with polynomial complexity. In this perception, Khachian succeeded in 1979 to introduce a new ellipsoid method from approaches applied originally to nonlinear optimization.

Interior point methods outperform the Simplex ones, and they have recently been the subject of several monographs including Bonnans and Gilbert [5], Evtushenko and Zhadan [6], Nesterov and Emirovski [7], and Wright [8] and Ye [9].

Interior point methods can be classified into three different groups as follows: projective methods and their alternatives as in Powell [10] and Rosen [11,12], central trajectory methods (see Ouriemchi [13] and Forsgren et al. [14]), and barrier/penalty methods, where majorant functions were originally proposed by Crouzeix and Merikhi [15] to solve a semidefinite optimization problem. Inspired by this work, Menniche and Benterki [16] and Bachir Cherif and Merikhi [17] applied this idea to linear and nonlinear optimizations, respectively.



Citation: Leulmi, A. An Efficient Penalty Method without a Line Search for Nonlinear Optimization. *Axioms* 2024, 13, 176. https://doi.org/ 10.3390/axioms13030176

Academic Editors: Milena J. Petrović, Predrag S. Stanimirovic and Gradimir V. Milovanović

Received: 18 January 2024 Revised: 1 March 2024 Accepted: 3 March 2024 Published: 7 March 2024



**Copyright:** © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). A majorant function for the penalty method in convex quadratic optimization was proposed by Chaghoub and Benterki [18]. On the other hand, A. Leulmi et al. [19,20] used new minorant functions for semidefinite optimization, and this idea was extended to linear programming by A. Leulmi and S. Leulmi in [21].

As far as we know, our new approximate function has not been studied in the nonlinear optimization literature. These approximate functions are more convenient and efficient than the line search method for rapidly computing the displacement step.

Therefore, in our work, we aim to optimize a nonlinear problem based on prior efforts. Thus, we propose a straightforward and effective barrier penalty method using new minorant functions.

More precisely, we first introduce the position of the problem and its perturbed problem with the results of convergence in Sections 2 and 3 of our paper. Then, in Section 4, we establish the solution of the perturbed problem by finding new minorant functions. Section 5 is devoted to presenting a concise description of the algorithm and to illustrating the outperformance of our new approach by carrying out a simulation study. Finally, we summarize our work in the conclusion.

Throughout this paper, the following notations are adopted. Let  $\langle ., . \rangle$  and  $\|.\|$  denote the scalar product and the Euclidean norm, respectively, given by the following:

$$\langle x, y \rangle = x^T y = \sum_{i=1}^n x_i y_i, \ x, y \in \mathbb{R}^n$$

and

$$||x|| = \sqrt{\langle x, x \rangle} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

## 2. The Problem

We aim to present an algorithm for solving the following optimization problem:

$$\min_{\substack{Ax = b \\ x \ge 0,}} f(x)$$
 (P)

where  $b \in \mathbb{R}^m$  and  $A \in \mathbb{R}^{m \times n}$  is a full-rank matrix with m < n. For this purpose, we need the following hypothesis:

**Hypothesis 1.** *f* is nonlinear, twice continuously differentiable, and convex on  $\mathcal{L}$ , where  $\mathcal{L} = \{x \in \mathbb{R}^n : Ax = b; x \ge 0\}$  is the set of realizable solutions of (P).

**Hypothesis 2.** (*P*) satisfies the condition of interior point (IPC), i.e., there exists  $x_0 > 0$  such that  $Ax_0 = b$ .

**Hypothesis 3.** *The set of optimal solutions of (P) is nonempty and bounded.* 

Notice that these conditions are standard in this context. We refer to [17,20].

If  $x^*$  is an optimal solution, there exist two Lagrange multipliers  $p^* \in \mathbb{R}^m$  and  $q^* \in \mathbb{R}^n$ , such that

$$\begin{cases} \nabla f(x^*) + A^T p^* = q^* \ge 0, \\ Ax^* = b, \\ \langle q^*, x^* \rangle = 0. \end{cases}$$
(1)

#### 3 of 14

## 3. Formulation of the Perturbed Problem of (P)

Let us first consider the function  $\psi$  defined on  $\mathbb{R} \times \mathbb{R}^n$  by the following:

$$\psi(\eta, x) = \begin{cases} f(x) + \sum_{i=1}^{n} \xi(\eta, x_i) & \text{if} \\ +\infty & \text{if not,} \end{cases} x \ge 0, \ Ax = b$$

where  $\xi : \mathbb{R}^2 \longrightarrow (-\infty, +\infty]$  is a convex, lower semicontinuous and proper function given by the following:

$$\xi(\eta, \alpha) = \begin{cases} \eta \ln(\eta) - \eta \ln(\alpha) & \text{if} & \alpha > 0 \text{ and } \eta > 0, \\ 0 & \text{if} & \alpha \ge 0 \text{ and } \eta = 0, \\ +\infty & \text{otherwise.} \end{cases}$$

Thus,  $\psi$  is a proper, convex, and lower semicontinuous function. Furthermore, the function *g* defined by

$$g(\eta) = \inf_{x \in \mathbb{R}^n} \left[ \psi_{\eta}(x) = f(x) + \sum_{i=1}^n \xi(\eta, x_i) \right]$$
(P $\eta$ )

is convex. Notice that for  $\eta = 0$ , the perturbed problem  $(P\eta)$  coincides with the initial problem (P); then,  $f^* = g(0)$ .

## 3.1. Existence and Uniqueness of Optimal Solution

To show that the perturbed problem  $(P\eta)$  has a unique optimal solution, it is sufficient to demonstrate that the recession cone of  $\psi_{\eta}$  is reduced to zero.

**Proof.** For a fixed  $\eta$ , the function  $\psi_{\eta}$  is proper, convex, and lower semicontinuous. The asymptotic function of  $\psi_{\eta}$  is defined by the following:

$$(\psi_{\eta})_{\infty}(d) = \lim_{\alpha \to +\infty} \frac{\psi_{\eta}(x_0 + \alpha d) - \psi_{\eta}(x_0)}{\alpha},$$

thus, the asymptotic functions of *f* and  $\psi_{\eta}$  satisfy the relation:

$$\left(\psi_{\eta}\right)_{\infty}(d) = \begin{cases} (f)_{\infty}(x) & \text{if} \quad d \ge 0, \ Ad = 0, \\ +\infty & \text{if not.} \end{cases}$$

Moreover, hypothesis H3 is equivalent to

$$\{d \in \mathbb{R}^n : (f)_{\infty}(x) \le 0, d \ge 0, Ad = 0\} = \{0\}.$$

Then,

$$\left\{d \in \mathbb{R}^n : \left(\psi_\eta\right)_\infty(d) \le 0\right\} = \{0\}$$

and from [17], for each non-negative real number  $\eta$ , the strictly convex problem  $(P\eta)$  admits a unique optimal solution noted by  $x_{\eta}^*$ . The solution of the problem (P) is the limit of the solutions sequence of the perturbed problem  $(P\eta)$  when  $\eta$  tends to 0.  $\Box$ 

## 3.2. Convergence of the Solution

Now, we are in a position to state the convergence result of  $(P\eta)$  to (P), which is proved in Lemma 1 on [18].

Let  $\eta > 0$ , for all  $x \in \mathcal{L}$ ; we define  $\psi(x, \eta) = f_{\eta}(x)$ .

**Lemma 1** ([18]). We consider  $\eta > 0$ . If the perturbed problem  $(P\eta)$  admits an optimal solution  $x_{\eta}$ , such that  $\lim_{n \to \infty} x_{\eta} = x^*$ , then the problem (P) admits an optimal solution  $x^*$ .

We use the classical prototype of penalty methods. We begin our process with  $(x_0, \eta_0) \in \tilde{\mathcal{L}} \times (0, \infty)$ , where

$$\tilde{\mathcal{L}} = \{ x \in \mathbb{R}^n : x > 0, \ Ax = b \}$$
(2)

and the iteration scheme is divided into the following steps:

1. Select  $\eta_{k+1} \in (0, \eta_k)$ .

2. Establish an approximate solution  $x_{k+1}$  for  $(P\eta_k)$ . It is obvious that  $\psi(\eta_k, x_{k+1}) < \psi(\eta_k, x_k)$ .

**Remark 1.** If the values of the objective functions of the problem (P) and the perturbed problem  $(P\eta)$  are equal and finite, then (P) will have an optimal solution if and only if  $(P\eta)$  has an optimal solution.

The iterative process stops when we obtain an acceptable approximation of g(0).

## 4. Computational Resolution of the Perturbed Problem

Our approach to the numerical solution of the perturbed problem  $(P\eta)$ , consists of two stages. In the first one, we calculate the descent direction using the Newton approach, and in the second one, we propose an efficient new-minorant-functions approach to compute the displacement step easily and quickly relative to the line search method.

## 4.1. The Descent Direction

As  $(P\eta)$  is strictly convex, the necessary and sufficient optimality conditions state that  $x_{\eta}$  is an optimal solution of  $(P\eta)$  if and only if it satisfies the nonlinear system:

$$\nabla \psi_{\eta}(x_{\eta}) = 0.$$

Using the Newton approach, a penalty method is provided to solve the above system, where the vector  $x_{k+1}$  in each is given by  $x_{k+1} = x_k + \alpha_k d_k$ .

The solution of the following quadratic convex optimization problem is necessary to obtain the Newton descent direction d:

$$\min_{d} \left[ \left\langle \nabla \psi_{\eta}(x), d \right\rangle + \frac{1}{2} \left\langle \nabla^{2} \psi_{\eta}(x) d, d \right\rangle : Ad = 0 \right] = \min_{d} [H(\eta, x) : Ad = 0],$$

where  $x \in \tilde{\mathcal{L}}$  and

$$\begin{split} \psi_{\eta}(x) &= f(x_{\eta}) + n\eta \ln \eta - \eta \sum_{i=1}^{n} \ln(x_{i}) \\ \nabla \psi_{\eta}(x) &= \nabla f(x) - \eta X^{-1}e \\ \nabla^{2}\psi_{\eta}(x) &= \nabla^{2}f(x) + \eta X^{-2}e \\ H(\eta, x) &= \langle \nabla \psi_{\eta}(x), d \rangle + \frac{1}{2} \langle \nabla^{2}\psi_{\eta}(x)d, d \rangle, \end{split}$$

with the diagonal matrix  $X = \operatorname{diag}(x_i)_{i=\overline{1,n}}$ .

The Lagrangian is given by the following:

$$L(x,s) = \left\langle \nabla f(x) - X^{-1}\eta, d \right\rangle + \frac{1}{2} \left\langle \left( \nabla^2 f(x) + X^{-2}\eta \right) d, d \right\rangle + \left\langle Ad, s \right\rangle,$$

where  $s \in \mathbb{R}^m$  is the Lagrange multiplier. It is sufficient for solving the linear system equations with n + m:

$$\begin{cases} \nabla f(x) - \eta X^{-1}e + \langle \left(\nabla^2 f(x) + X^{-2}\eta\right)d, d \rangle + A^t s = 0\\ Ad = 0, \end{cases}$$

then,

$$\begin{aligned} & \left(\nabla^2 f(x) - \eta X^{-2}\right) d + A^T s &= \eta X^{-1} e - \nabla f(x) \\ & A d &= 0. \end{aligned}$$

It is simple to prove that system (3) is non-singular. We obtain

$$\begin{cases} d^T (\nabla^2 f(x) - \eta X^{-2}) d + d^T A^T s = d^T \eta X^{-1} e - d^T \nabla f(x) \\ d^T A d = 0, \end{cases}$$

As  $d^T A^T s = (Ad)^T s = 0$  and Ad = 0, we obtain

$$\left\langle \nabla^2 f(x)d,d\right\rangle + \left\langle \nabla f(x),d\right\rangle = \eta \left[\left\langle X^{-1}d,e\right\rangle - \left\|X^{-1}d\right\|^2\right].$$
 (4)

The system can also be written as follows:

$$\begin{cases} \left(X\nabla^2 f(x)X\right)\left(X^{-1}d\right) + \eta I\left(X^{-1}d\right) + XA^T s = \eta XX^{-1}e - X\nabla f(x) \\ AX(X^{-1}d) = 0, \end{cases}$$
(5)

Thus, the Newton descent direction is obtained. Throughout this paper, we take *x* instead of  $x_{\eta}$ .

#### 4.2. Computation of the Displacement Step

This section deals with the numerical solution of the displacement step. We give a brief highlight of the line search methods used in nonlinear optimization problems. Then, we collect some important results of approximate function approaches applied to both semidefinite and linear programming problems. Finally, we propose our new approximate function method for the nonlinear optimization problem (P).

#### 4.2.1. Line Search Methods

The line search methods consists of determining a displacement step  $\alpha_k$ , which ensures the sufficient decrease in the objective at each iteration  $x_{k+1} = x_k + \alpha_k d_k$ , where  $\alpha_k > 0$ , along the descent direction  $d_k$ ; in other words, it involves solving the following onedimensional problem:

$$\varphi(\alpha) = \min_{\alpha > 0} \psi_{\eta}(x_k + \alpha d_k).$$

The disadvantage of this method is that the solution  $\alpha$  is not necessarily optimal, which make the feasibility of  $x_{k+1}$  not guaranteed.

The line search techniques of Wolfe, Goldstein-Armijo, and Fibonacci are the most widely used ones. However, generally, their computational volume is costly. This is what made us search for another alternative.

## 4.2.2. Approximate Functions Techniques

These methods are based on sophisticated techniques introduced by J.P. Crouzeix et al. [15] and A. Leulmi et al. [20] to obtain the solution of a semidefinite optimization problem.

The aim of these techniques is to give a minimized approximation of one real-variable function  $\varphi(\alpha)$  defined by

$$\begin{split} \varphi(\alpha) &= \frac{1}{\eta} [\psi_{\eta}(x + \alpha d) - \psi_{\eta}(x)] \\ &= \frac{1}{\eta} [f(x + \alpha d) - f(x)] - \sum_{i=1}^{n} \ln(1 + \alpha t_{i}), \ t = X^{-1} d. \end{split}$$

The function  $\varphi$  is convex, and we obtain the following:

$$\begin{split} \varphi'(\alpha) &= \frac{1}{\eta} \langle \nabla f(x+\alpha d), d \rangle - \sum_{i=1}^{n} \frac{t_i}{1+\alpha t_i}, \\ \varphi''(\alpha) &= \frac{1}{\eta} \langle \nabla^2 f(x+\alpha d), d \rangle + \sum_{i=1}^{n} \frac{t_i^2}{(1+\alpha t_i)^2}. \end{split}$$

We find that  $\varphi'(0) + \varphi''(0) = 0$ , deduced from (4), which is expected since *d* is the direction of Newton's descent direction.

We aim to avoid the disadvantages of line search methods and accelerate the convergence of the algorithm. For this reason, we have to identify an  $\bar{\alpha}$  that yields a significant decrease in the function  $\varphi(\alpha)$ . This is the same as solving a polynomial equation of degree n + 1, where f is a linear function.

Now, we include a few helpful inequalities below, which are used throughout the paper.

H. Wolkowicz et al. [22] see also Crouzeix and Seeger [23] presented the following inequalities:

$$\begin{aligned} \overline{z} &- \sigma_z \sqrt{n-1} &\leq & \min_i z_i \leq \overline{z} - \frac{\sigma_z}{\sqrt{n-1}} \\ \overline{z} &+ \frac{\sigma_x}{\sqrt{n-1}} &\leq & \max_i z_i \leq \overline{z} + \sigma_z \sqrt{n-1}, \end{aligned}$$

where  $\overline{z}$  and  $\sigma_z$  represent the mean and the standard deviation, respectively, of a statistical real numbers series  $\{z_1, z_2, ..., z_n\}$ . The later quantities are defined as follows:

$$\overline{z} = \frac{1}{n} \sum_{i=1}^{n} z_i$$
 and  $\sigma_z^2 = \frac{1}{n} \sum_{i=1}^{n} z_i^2 - \overline{z}^2 = \frac{1}{n} \sum_{i=1}^{n} (z_i - \overline{z})^2.$ 

**Theorem 1** ([15]). *Let*  $z_i > 0$ , *for* i = 1, 2, ..., n. *We have the following:* 

$$\sum_{i=1}^{n} \ln(z_i) \le \ln\left(\overline{z} + \sigma_z \sqrt{n-1}\right) + (n-1)\ln\left(\overline{z} - \frac{\sigma_z}{\sqrt{n-1}}\right).$$
(6)

where  $z_i = 1 + \alpha t_i$ ,  $z = 1 + \alpha t$ , and  $\sigma_z = \alpha \sigma_t$ .

We will proceed to present the paper's principal result.

4.2.3. New Approximate Functions Approach

Let

$$\varphi(\alpha) = \frac{1}{\eta} [f(x + \alpha d) - f(x)] - \sum_{i=1}^{n} \ln(1 + \alpha t_i),$$

be defined on  $\tilde{\alpha} = \min_{i \in I_-} \left\{ \frac{-1}{t_i} \right\}$  such that  $I_- = \{i : t_i < 0\}$ .

To find the displacement step, it is necessary to solve  $\varphi'(\alpha) = 0$ . Considering the difficulty of solving a non-algebraic equation, approximate functions are recommended alternatives.

Two novel approximation functions of  $\varphi$  are introduced in the following lemma.

**Lemma 2.** For all  $\alpha \in [0, \alpha_1^*[$  with  $\alpha_1^* = \min(\widehat{\alpha}, \widehat{\alpha}_1)$ , we have

$$\varphi(\alpha) \geq \widehat{\varphi}_1(\alpha)$$

and for all  $\alpha \in [0, \alpha_2^*[$  with  $\alpha_2^* = \min(\widehat{\alpha}, \widehat{\alpha}_2)$ , we obtain

$$\widehat{\varphi}_2(\alpha) \leq \varphi(\alpha),$$

where

$$\widehat{\varphi}_1(\alpha) = \frac{1}{\eta} (f(x + \alpha d) - f(x)) - \ln(1 + \delta \alpha) - (n-1)\ln(1 + \beta \alpha),$$

and

with

$$\widehat{\varphi}_2(\alpha) = \frac{1}{\eta} (f(x + \alpha d) - f(x)) - \tau \ln(1 + \beta_1 \alpha),$$

 $\begin{cases} \beta = \overline{t} - \frac{\sigma_t}{\sqrt{n-1}} \\ \delta = \overline{t} + \sigma_t \sqrt{n-1} \\ \beta_1 = \frac{\|t\|^2}{n\overline{t}}. \end{cases}$ (7)

Furthermore, we have

$$\widehat{\varphi}_2(\alpha) \leq \widehat{\varphi}_1(\alpha) \leq \varphi(\alpha),$$

**Proof.** We start by proving that

$$\varphi(\alpha) \geq \widehat{\varphi}_1(\alpha),$$

Theorem 1 gives

$$\sum_{i=1}^{n} \ln(z_i) \le \ln\left(\overline{z} + \sigma_z \sqrt{n-1}\right) + (n-1)\ln\left(\overline{z} - \frac{\sigma_z}{\sqrt{n-1}}\right),$$

then,

$$\sum_{i=1}^{n} \ln(1+\alpha t_i) \le \ln(1+\alpha\delta) + (n-1)\ln(1+\alpha\beta)$$

and

$$-\sum_{i=1}^{n}\ln(1+\alpha t_i) \geq -\ln(1+\alpha\delta) - (n-1)\ln(1+\alpha\beta).$$

Hence,

$$\frac{1}{\eta}(f(x+\alpha d) - f(x)) - \sum_{i=1}^{n} \ln(1+\alpha t_i) \ge \frac{1}{\eta}(f(x+\alpha d) - f(x)) - \ln(1+\alpha\delta) - (n-1)\ln(1+\alpha\beta).$$

Therefore,

$$\begin{split} \varphi(\alpha) &\geq \quad \widehat{\varphi}_1(\alpha) = \frac{1}{\eta} (f(x+\alpha d) - f(x)) - \ln(1+\alpha \delta) - (n-1)\ln(1+\alpha \beta), \\ \widehat{\varphi}_1'(\alpha) &= \quad \frac{1}{\eta} \langle \nabla f(x+\alpha d) - d \rangle - \frac{\delta}{1+\alpha \delta} - (n-1)\frac{\beta}{1+\alpha \beta}. \end{split}$$

Let us consider the following:

$$g(\alpha) = \varphi(\alpha) - \widehat{\varphi}_2(\alpha).$$

We have

$$g''(\alpha) = \sum_{i=1}^{n} \frac{t_i^2}{(1+\alpha t_i)^2} - \tau \frac{\beta_1}{(1+\alpha \beta_1)^2}.$$

Because of the fact that  $|t_i| \le ||t||$  and  $n\overline{t} \le ||t||$ , it is easy to see that  $\forall \alpha \ge 0$ ,  $g''(\alpha) \ge 0$ . Therefore,

$$\frac{1}{\eta}(f(x+\alpha d) - f(x)) - \sum_{i=1}^{n} \ln(1+\alpha t_i) \ge \frac{1}{\eta}(f(x+\alpha d) - f(x)) - \tau \ln(1+\beta_1 \alpha),$$

then,  $\varphi(\alpha) \leq \widehat{\varphi}_2(\alpha)$ .  $\Box$ 

Hence, the domain of  $(\hat{\varphi}_i)_{i=1,2}$  is included in the domain of  $\varphi$ , which is  $(0, \tilde{\alpha})$ , where

$$\widetilde{\alpha} = \max[\alpha : 1 + \alpha \delta > 0, \ 1 + \alpha \beta > 0].$$

Let us remark that

$$0 = \widehat{\varphi}_i(0) = \varphi(0), \ -\varphi'(0) = -\widehat{\varphi}'_i(0) = \varphi''_i(0) = \widehat{\varphi}''_i(0) > 0, \ i = 1, 2$$

Thus,  $\varphi$  is well approximated by  $\hat{\varphi}_i$  in a neighborhood of 0. Since  $\hat{\varphi}_i$  is strictly convex, it attains its minimum at one unique point  $\bar{\alpha}$ , which is the unique root of the equation  $\hat{\varphi}'_i(\alpha) = 0$ . This point belongs to the domain of  $\hat{\varphi}_i$  (i = 1, 2). Therefore,  $\varphi$  is bounded from below by  $\hat{\varphi}_1$ :

$$\widehat{\varphi}_1(\bar{\alpha}) \leq \varphi(\bar{\alpha}) < 0$$

And it is also bounded from below by  $\hat{\varphi}_2$ :

$$\widehat{\varphi}_2(\bar{\alpha}) \leq \varphi(\bar{\alpha}) < 0$$

Then,  $\bar{\alpha}$  gives an apparent decrease in the function  $\varphi$ .

4.3. Minimize an Auxiliary Function

We now consider the minimization of the function

$$\varphi_1(\alpha) = n\gamma\alpha - \ln(1+\delta\alpha) - (n-1)\ln(1+\beta\alpha),$$

and we also have the following approximate function:

$$\varphi_2(\alpha) = n\gamma\alpha - \tau \ln(1 + \beta_1 \alpha),$$

where  $\beta_1$  is defined in (7). Then, we have the following:

$$\varphi_1'(\alpha) = n\gamma - \frac{\delta}{1+\delta\alpha} - (n-1)\frac{\beta}{1+\beta\alpha},$$
  
$$\varphi_1''(\alpha) = \frac{\delta^2}{(1+\delta\alpha)^2} + (n-1)\frac{\beta^2}{(1+\beta\alpha)^2}.$$

and

$$\begin{aligned} \varphi_2'(\alpha) &= n\gamma - \tau \frac{\beta_1}{(1+\beta_1\alpha)}, \\ \varphi_2''(\alpha) &= \tau \frac{\beta_1^2}{(1+\beta_1\alpha)^2}. \end{aligned}$$

We remark that for i = 1, 2:

$$\varphi_i(0) = 0, \ \varphi_i'(0) = n(\gamma - \overline{t}), \ \varphi_i''(0) = ||t||^2.$$

We present the conditions  $\varphi'_i(0) < 0$  and  $\varphi''_i(0) > 0$ . The function  $\varphi$  is strictly convex. It attains its minimum at one unique point  $\alpha$  such that  $\varphi'_i(\alpha) = 0$ , which is one of the roots of the equations

$$\gamma\delta\beta\alpha^2 + \alpha(\gamma\delta + \gamma\beta - \delta\beta) + \gamma - \bar{t} = 0, \tag{8}$$

and

$$n\gamma\beta(1+\beta\alpha) - \|t\|^2 = 0.$$
<sup>(9)</sup>

For Equation (8), the roots are explicitly calculated, and we distinguish the following cases:

• If  $\delta = 0$ , we obtain  $\bar{\alpha}_1 = \frac{\bar{t} - \gamma}{\gamma \beta}$ .

- If  $\beta = 0$ , we obtain  $\bar{\alpha}_1 = \frac{\bar{t} \gamma}{\gamma \delta}$ .
- If  $\gamma = 0$ , we have  $\bar{\alpha}_1 = \frac{-\bar{t}}{\delta\beta}$ .
- If  $\gamma \delta \beta \neq 0$ ,  $\bar{\alpha}$  is the only root of the second-degree equation that belongs to the domain of definition of  $\varphi$ . We obtain  $\triangle = \frac{1}{\gamma^2} + \frac{1}{\beta^2} + \frac{1}{\delta^2} \frac{2}{\delta\beta} + \left(\frac{2n-4}{n}\right) \left[\frac{1}{\beta\gamma} \frac{1}{\gamma\delta}\right]$ . Both roots are

$$ar{lpha}_{1.1} = rac{1}{2} igg( rac{1}{\gamma} - rac{1}{eta} - rac{1}{\delta} + \sqrt{ riangle} igg),$$

and

$$\bar{\alpha}_{1.2} = \frac{1}{2} \left( \frac{1}{\gamma} - \frac{1}{\beta} - \frac{1}{\delta} + \sqrt{\bigtriangleup} \right)$$

Then, the root of Equation (9) is explicitly calculated, and we have

$$\bar{\alpha}_2 = \left( \frac{\|t\|^2}{\beta(1-\beta)} - \frac{1}{\beta} 
ight).$$

Consequently, we compute the two values  $\bar{\alpha}_i$ , i = 1, 2, explicitly. Then, we take  $\bar{\alpha}_1, \bar{\alpha}_2 \in [0, \tilde{\alpha} - \varepsilon[$ , where  $\varepsilon > 0$  is a fixed precision and  $\varphi'(\bar{\alpha}_i) > 0$ , i = 1, 2.

**Remark 2.** The computation of  $\bar{\alpha}_i$ , i = 1, 2 is performed through a dichotomous procedure in the cases where  $\bar{\alpha}_i \notin (0, \tilde{\alpha} - \varepsilon)$ , and  $\varphi'(\bar{\alpha}_i) > 0$ , as follows:

Put a = 0,  $b = \tilde{\alpha} - \varepsilon$ . While  $|b - a| > \varepsilon$  do If  $\varphi'(\frac{a+b}{2}) < 0$  then,  $b = \frac{a+b}{2}$ , else  $a = \frac{a+b}{2}$ , so  $\bar{\alpha}_i = b$ .

This computation guarantees a better approximation of the minimum of  $\varphi'(\alpha)$  while remaining in the domain of  $\varphi$ .

#### 4.4. The Objective Function f Is

4.4.1. Linear

For all *x*, there exists  $c \in \mathbb{R}^n$  such that  $f(x) = \langle c, x \rangle$ . The minimum of  $\hat{\varphi}_i$  is reached at the unique root  $\bar{\alpha}$  of the equation  $\varphi'(\alpha) = 0$ . Then,

$$\widehat{\varphi}_i(\bar{\alpha}) \le \varphi(\bar{\alpha}) < \widehat{\varphi}_i(0) = \varphi(0) = 0$$

Take  $\gamma = n^{-1} \langle c, d \rangle$  in the auxiliary function  $\varphi$ . The two functions  $\varphi$  and  $(\widehat{\varphi}_i)_{i=1,2}$  coincide.

 $\bar{\alpha}$  yields a significant decrease in the function  $\psi_{\eta}$  along the descent direction *d*. It is interesting to note that the condition  $\hat{\varphi}'_i(0) + \hat{\varphi}''_i(0) = 0$  (*i* = 1, 2) implies the following:

$$-\widehat{\varphi}_i'(0) = n(\overline{t} - \gamma) = ||t||^2 = n(\overline{t}^2 - \sigma_t^2) = \widehat{\varphi}_i''(0) > 0.$$

4.4.2. Convex

 $\nabla f(x + \alpha d)$  is no longer constant, and the equation  $\widehat{\varphi}'_i(\alpha) = 0$  is not reduced to one equation of a second degree for i = 1, 2.

We consider another function  $\tilde{\varphi}$  less than  $\varphi$ . Given  $\hat{\alpha} \in (0, \tilde{\alpha})$ , we have, for all  $\alpha \in (0, \hat{\alpha}]$ , the following:

$$\frac{f(x+\alpha d)-f(x)}{\eta} \le \frac{f(x+\widehat{\alpha} d)-f(x)}{\eta\widehat{\alpha}}\alpha,$$
(10)

then,

$$\varphi(\alpha) \geq \widetilde{\varphi}_1(\alpha) = \frac{f(x + \widehat{\alpha}d) - f(x)}{\eta \widehat{\alpha}} \alpha - \ln(1 + \alpha\delta) - (n - 1)\ln(1 + \alpha\beta)$$

and

$$\varphi(\alpha) \geq \widetilde{\varphi}_2(\alpha) = \frac{f(x + \widehat{\alpha}d) - f(x)}{\eta \widehat{\alpha}} \alpha - \tau \ln(1 + \beta_1 \alpha), \ \tau \in ]0,1[.$$

We choose  $\gamma = \frac{f(x+\hat{\alpha}d)-f(x)}{n\eta\hat{\alpha}}$  in the auxiliary function  $\varphi$ , and we compute the root  $\bar{\alpha}$  of the equation  $\varphi'_i(\alpha) = 0$  with i = 1, 2.

Therefore, we have two cases:

1. Where  $\bar{\alpha} \leq \hat{\alpha}$ : We have the following:

$$\varphi(\bar{\alpha}) \geq \widehat{\varphi}_i(\bar{\alpha}) \geq \widetilde{\varphi}_i(\bar{\alpha}), \text{ for } i = 1, 2$$

and, thus, along the direction d, we obtain a significant decrease in the function  $\psi_{\eta}$ . The approximation accuracy of  $\varphi$  by  $\tilde{\varphi}_i$  being better for small values of  $\hat{\alpha}$  (for i = 1, 2), it is recommended to use a new value of  $\hat{\alpha}$ , situated between  $\tilde{\alpha}$  and the former  $\hat{\alpha}$ , for the next iteration. Moreover, the cost of the supplementary computation is small since it is the cost of one evaluation of f and the resolution of a second-order equation.

2. Where  $\bar{\alpha} > \hat{\alpha}$ : The computation of  $\hat{\alpha}$  is performed through a dichotomous procedure (see Remark 3).

## 5. Description of the Algorithm and Numerical Simulations

5.1. Description of the Algorithm

This section is devoted to introducing our algorithm for obtaining an optimal solution  $\bar{x}$  of (*P*).

Begin

Initialization  $\varepsilon > 0$  is a given precision.  $\hat{\eta} > 0$  and  $\sigma \in [0, 1]$  are given.  $x_0$  is a strictly realizable solution from  $\tilde{\mathcal{L}}, d_0 \in \mathbb{R}^m$ . Iteration

- 1. Start with  $\eta > \hat{\eta}$ .
- 2. Calculate *d* and  $t = X^{-1}d$ .
- 3. If  $||t|| > \varepsilon$ , calculate  $\overline{t}$ ,  $\gamma$ ,  $\delta$ ,  $\beta$ , and  $\beta_1$ .
- 4. Determine  $\bar{\alpha}$  following (8), (10), or (9) depending on the linear or nonlinear case.
- 5. Take the new iterate  $x = x + \bar{\alpha}d = X(e + \bar{\alpha}t)$  and go back to step 2.
- 6. If  $||t|| \le \varepsilon$ , a well approximation of  $g(\eta)$  has been obtained.
  - (a) If  $\eta \ge \hat{\eta}$  and  $\eta = \sigma \eta$ , return to step 2.
  - **(b)** If  $\eta < \hat{\eta}$ , STOP: a well approximate solution of (*P*) has been obtained.

End algorithm.

The aim of this method is to reduce the number of iterations and the time consumption. In the next section, we provide some examples.

## 5.2. Numerical Simulations

To assess the superior performance and accuracy of our algorithm, based on our minorant functions, numerical tests are conducted to make comparisons between our new approach and the classical line search method.

For this purpose, in this section, we present comparative numerical tests on different examples taken from the literature [5,24].

We report the results obtained by implementing the algorithm in MATLAB on an Intel Core i7-7700HQ (2.80 GHz) machine with 16.00 Go RAM.

5.2.1. Examples with a Fixed SizeNonlinear Convex ObjectiveExample 1. Let us take the following problem:

$$\begin{cases} \min 2x_1^2 + 2x_2^2 - 2x_1x_2 - 4x_1 - 6x_2 \\ x_1 + x_2 + x_3 = 2 \\ x_1 + 5x_2 + x_4 = 5 \\ x_1, x_2, x_3, x_4 \ge 0. \end{cases}$$

The optimal value is -7.1613, and the optimal solution is  $x^* = (1.1290 \quad 0.7742 \quad 0.0968 \quad 0)^t$ .

**Example 2.** *Let us take the following problem:* 

 $\begin{cases} \min x_1^3 + x_2^3 \\ x_1 - x_2 + x_3 + x_4 = 3 \\ 2x_1 + x_2 - x_3 + x_4 = 2.0086 \\ x_1 + x_3 + 2x_4 = 4.9957 \\ x_1, x_2, x_3, x_4 \ge 0. \end{cases}$ 

*The optimal value is* 0.0390, *and the optimal solution is*  $x^* = (0.3391 \ 0 \ 0.6652 \ 1.9957)^t$ .

**Example 3.** Let us consider the following problem:

$$\begin{cases} \min x_1^3 + x_2^3 + x_1x_2 \\ 2x_1 - x_2 + x_3 = 8 \\ x_1 + 2x_2 + x_4 = 6 \\ x_1, x_2, x_3, x_4 \ge 0. \end{cases}$$

*The optimal value is* 1.6157, *and the optimal solution is*  $x^* = (1.1734 \ 0 \ 5.6532 \ 4.8265)^t$ .

Example	st1		st2		LS	
	iter	Time (s)	iter	Time (s)	iter	Time (s)
1	12	0.0006	19	0.0015	6	0.0091
2	5	0.0004	9	0.0009	44	0.099
3	3	0.0001	5	0.0006	65	0.89

This table presents the results of the previous examples:

5.2.2. Example with a Variable Size The Objective Function *f* Is

**1-Linear:** Let us consider the linear programming problem:

$$\zeta = \min[c^T x : x \ge 0, \ Ax = b],$$

where *A* is an  $m \times 2m$  matrix given by the following:

$$A[i,j] = \begin{cases} 1 & \text{if } i = j & \text{or } j = i + m \\ 0 & \text{if not,} \end{cases}$$
  
$$c[i] = -1, c[i+m] = 0 \text{ and } b[i] = 2, \forall i = 1, \dots m,$$

where  $c, b \in \mathbb{R}^{2m}$ .

The results are presented in the table below.

Size	st1		st2		LS	
	iter	Time (s)	iter	Time (s)	iter	Time (s)
$5 \times 10$	1	0.0021	2	0.0039	9	0.0512
20  imes 40	1	0.0031	3	0.0045	13	0.0821
50  imes 100	2	0.0049	3	0.0032	17	0.3219
100  imes 200	2	0.0053	4	0.0088	19	0.5383
200  imes 400	2	0.0088	4	0.0098	22	0.9220
250  imes 500	3	0.0096	5	0.0125	26	9.2647

## 2-Nonlinear:

**Example 4** (Quadratic case [13]). *Let the quadratic problem be as follows:* 

$$\zeta = \min[f(x) : x \ge 0, \ Ax \ge b],$$

with  $f(x) = \frac{1}{2} \langle x, Qx \rangle$ , Q is the matrix defined for n = 2m by the following:

$$Q[i, j] = \begin{cases} 2j - 1 & \text{if } i > j \\ 2i - 1 & \text{if } i < j \\ i(i + 1) - 1 & \text{if } i = j, i, j = 1, ..., n \end{cases}$$
  

$$A[i, j] = \begin{cases} 1 & \text{if } i = j \text{ or } j = i + m, i = 1, ..., m \text{ and } j = 1, ..., n \\ 0 & \text{if not} \end{cases}$$
  

$$c[i] = -1, c[i + m] = 0 \text{ and } b[i] = 2, \forall i = 1, ..., m.$$

*This example is tested for many values of n.* 

The obtained results are given by the following table:

ex( <i>m</i> , <i>n</i> )	st1		st2		LS	
	iter	Time (s)	iter	Time (s)	iter	Time (s)
$300 \times 600$	5	0.9968	4	0.9699	26	19.5241
400 imes 800	7	18.1448	5	9.6012	35	86.1259
600  imes 1200	12	36.3259	5	19.0099	23	98.2354
1000  imes 2000	21	56.9912	17	41.1012	33	109.2553
$1500 \times 3000$	28	140.1325	23	95.6903	40	1599.1596

**Example 5** (The problem of Erikson [25]). *Let the following be the quadratic problem:* 

$$\zeta = \min\left[f(x) = \sum_{i=1}^n x_i \ln\left(\frac{x_i}{a_i}\right) : x_i + x_{i+m} = b, x \ge 0\right],$$

where n = 2m,  $a_i > 0$  and  $b \in \mathbb{R}^m$  are fixed.

This example is tested for different values of n,  $a_i$ , and  $b_i$ .

The following table results the obtained results in the case  $(a_i = 2, \forall i = 1, ..., n, b_i = 4, \forall i = 1, ..., m)$ :

ex( <i>m</i> , <i>n</i> )	st1		st2		LS	
	iter	Time (s)	iter	Time (s)	iter	Time (s)
10  imes 20	1	0.0001	2	0.0012	4	0.0236
40  imes 100	2	0.0021	3	0.0033	5	0.7996
100  imes 200	2	0.0043	3	0.0201	5	1.5289
500  imes 1000	2	3.0901	4	5.9619	12	22.1254

In the above tables, we take  $\varepsilon = 1.0 \times 10^{-4}$ .

We also denote the following:

- (iter) is the number of iterations.
- (time) is the computational time in seconds (s).

-  $(sti)_{i=1,2}$  represents the strategy of approximate functions introduced in this paper.

- (LS) represents the classical line search method.

**Commentary:** The numerical tests carried out show, without doubt, that our approach leads to a very significant reduction in the cost of calculation and an improvement in the result. When comparing the approximate functions to the line search approach, the number of iterations and computing time are significantly reduced.

## 6. Conclusions

The contribution of this paper is particular focused on the study of nonlinear optimization problems by using the logarithmic penalty method based on some new approximate functions. We first formulate the problems (*P*) and (*P* $\eta$ ) with the results of the convergence. Then, we find their solutions by using new approximate functions.

Finally, to lend further support to our theoretical results, a simulation study is conducted to illustrate the good accuracy of the studied approach. More precisely, our new approximate functions approach outperforms the line search one as it significantly reduces the cost and computing time.

Funding: This work has been supported by: The General Directorate of Scientific Research and Technological Development (DGRSDT-MESRS) under the PRFU project number C00L03UN190120220009. Algeria.

Data Availability Statement: Data are contained within the article.

Acknowledgments: The author is very pleased to thank the editor and the reviewers for their helpful suggestions and comments.

Conflicts of Interest: The authors declare no conflicts of interest.

#### References

- 1. Frank, M.; Wolfe, B. An algorithm for quadratic programming. Nav. Res. Logist. Q. 1956, 3, 95–110. [CrossRef]
- 2. Wolfe, P. A Duality Theorem for Nonlinear Programming. Q. Appl. Math. 1961, 19, 239–244. [CrossRef]
- 3. Bracken, J.; M1cCormiek, G.P. Selected Applications of Nonlinear Programming; John Wiley & Sons, Inc.: New York, NY, USA, 1968.
- 4. Fiacco, A.V.; McCormick, G.P. Nonlinear Programming: Sequential Unconstrained Minimization Techniques; John Wiley & Sons, Inc.: New York, NY, USA, 1968.
- Bonnans, J.-F.; Gilbert, J.-C.; Lemaréchal, C.; Sagastizàbal, C. Numerical Optimization: Theoretical and Practical Aspects; Mathematics and Applications; Springer: Berlin/Heidelberg, Germany, 2003; Volume 27.
- 6. Evtushenko, Y.G.; Zhadan, V.G. Stable barrier-projection and barrier-Newton methods in nonlinear programming. In *Optimization Methods and Software*; Taylor & Francis: Abingdon, UK, 1994; Volume 3, pp. 237–256.
- 7. Nestrov, Y.E.; Nemiroveskii, A. Interior-Point Polynomial Algorithms in Convex Programming; SIAM: Philadelphia, PA, USA, 1994.
- 8. Wright, S.J. *Primal–Dual Interior Point Methods*; SIAM: Philadelphia, PA, USA, 1997.
- Ye, Y. Interior Point Algorithms: Theory and Analysis. In *Discrete Mathematics Optimization*; Wiley-Interscience Series; John Wiley & Sons: New York, NY, USA, 1997.
- Powell, M.J.D. Karmarkar's Algorithm: A View from Nonlinear Programming; Department of Applied Mathematics and Theoretical Physics, University of Cambridge: Cambridge, UK, 1989; Volume 53.
- Rosen, J.B. The Gradient Projection Method for Nonlinear Programming. Soc. Ind. Appl. Math. J. Appl. Math. 1960, 8, 181–217. [CrossRef]

- 12. Rosen, J.B. The Gradient Projection Method for Nonlinear Programming. *Soc. Ind. Appl. Math. J. Appl. Math.* **1961**, *9*, 514–553. [CrossRef]
- 13. Ouriemchi, M. Résolution de Problèmes non Linéaires par les Méthodes de Points Intérieurs. Théorie et Algorithmes. Doctoral Thesis, Université du Havre, Havre, France, 2006.
- 14. Forsgren, A.; Gill, P.E.; Wright, M.H. Interior Methods for Nonlinear Optimization; SIAM: Philadelphia, PA, USA, 2002; Voluume 44, pp. 525–597.
- 15. Crouzeix, J.P.; Merikhi, B. A logarithm barrier method for semidefinite programming. *RAIRO-Oper. Res.* **2008**, *42*, 123–139. [CrossRef]
- 16. Menniche, L.; Benterki, D. A Logarithmic Barrier Approach for Linear Programming. J. Computat. Appl. Math. 2017, 312, 267–275. [CrossRef]
- 17. Cherif, L.B.; Merikhi, B. A Penalty Method for Nonlinear Programming. RAIRO-Oper. Res. 2019, 53, 29–38. [CrossRef]
- Chaghoub, S.; Benterki, D. A Logarithmic Barrier Method Based On a New Majorant Function for Convex Quadratic Programming. IAENG Int. J. Appl. Math. 2021, 51, pp 563-568.
- 19. Leulmi, A. Etude d'une Méthode Barrière Logarithmique via Minorants Functions pour la Programmation Semi-Définie. Doctoral Thesis, Université de Biskra, Biskra, Algeria, 2018.
- 20. Leulmi, A.; Merikhi, B.; Benterki, D. Study of a Logarithmic Barrier Approach for Linear Semidefinite Programming. J. Sib. Fed. Univ. Math. Phys. 2018, 11, 300–312.
- 21. Leulmi, A.; Leulmi, S. Logarithmic Barrier Method via Minorant Function for Linear Programming. J. Sib. Fed. Univ. Math. Phys. 2019, 12, 191–201. [CrossRef]
- 22. Wolkowicz, H.; Styan, G.P.H. Bounds for Eigenvalues Using Traces. Lin. Alg. Appl. 1980, 29, 471–506. [CrossRef]
- 23. Crouzeix, J.-P.; Seeger, A. New bounds for the extreme values of a finite sample of real numbers. *J. Math. Anal. Appl.* **1996**, 197, 411–426. [CrossRef]
- 24. Bazraa, M.S.; Sherali, H.D.; Shetty, C.M. Nonlinear Programming, Willey-Interscience; John Wiley & Sons, Inc.: Hoboken, NJ, USA; Toronto, ON, Canada, 2006.
- 25. Shannon, E. A mathematical theory of communication. Bell Syst. Tech. J. 1948, 27; 379–423+623–656. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.