LVI SBPO SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL
DE 04 ATÉ 07 NOVEMBRO 2024
FORTALEZA | CEARÁ

An Augmented Lagrangian Algorithm for Box-Constraint Optimization

Nelson Maculan
Programa de Engenharia de Sistemas e Computação
UFRJ/COPPE
maculan@cos.ufrj.br

Lennin Mallma Ramirez
Programa de Engenharia de Sistemas e Computação
UFRJ/COPPE
lenninmr@cos.ufrj.br

Igor Pereira dos Santos Pereira
Programa de Engenharia Química
UFRJ/COPPE
ipereira@peq.coppe.ufrj.br

ABSTRACT

In this work, we present a nonquadratic augmented Lagrangian algorithm called dislocation hyperbolic augmented Lagrangian algorithm (DHALA). This algorithm solves the nonconvex optimization problem with inequality and box constraints. Previously, this restriction of box constraints was not considered by this algorithm. Finally, we computationally solve some examples considering our proposed algorithm.

KEYWORDS. Augmented Lagrangian. Box constraints. Convergence.

Mathematical Programming

https://proceedings.science/p/193730?lang=en

1. Introduction

We are interested in the following nonconvex programming problem

$$\begin{aligned} (P) \qquad & \min \ f(x) \\ & s.t. \ \ g_i(x) \geq 0, \ \ i = 1, ..., m, \\ & x \in \Omega, \end{aligned}$$

where the set Ω is compact and convex whereas $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, ..., m$ are continuously differentiable on an open set that contains Ω. In general, $\Omega = \{x \in \mathbb{R}^n \ : \ l \leq x \leq u\}$. Classic works in the literature that solve the optimization problem with constraints are Hestenes [1969] and Powell [1969], where they solve the problem with equality constraints. Augmented Lagrangian algorithms now solve problems with inequality constraints, box constraints and nonconvexity assumptions, see Conn et al. [1997] and Birgin et al. [2005]. The algorithms mentioned are mainly based on the classical barrier and quadratic penalty function, respectively. In this work, we solve the problem (P) with the Dislocation Hyperbolic Augmented Lagrangian Algorithm (DHALA), that is the contribution of our work. The DHALA was originally proposed in Xavier [1992]. We also show computationally that by considering the box constraints in the subproblem generated by DHALA, our algorithm achieves convergence. The paper is organized as follows: Section 2, we present some basic results; in Section 3, we present our main convergence result; and in Section 4, we computationally illustrate our proposed algorithm.

2. Basic Results

The Lagrangian function for the problem (P) is

$$L(x, \lambda) = f(x) - \sum_{i=1}^{m} \lambda_i g_i(x), \tag{2.1}$$

where $\lambda_i \geq 0, \ \ i = 1, ..., m$, is the vector of Lagrange multipliers. In Xavier [1992] (also see Xavier [2001]), the dislocation hyperbolic penalty function (DHPF) is proposed and defined to solve problem (P) as follows:

$$l_H(x, \lambda, \tau) = f(x) - \sum_{i=1}^{m} \tau \ h\left(\frac{\lambda_i g_i(x)}{\tau}\right), \tag{2.2}$$

where the function $h : \mathbb{R} \to \mathbb{R}$ is defined as $h(t) = t - \sqrt{t^2 + 1} + 1$. Henceforth, we will call the h function as the dislocation hyperbolic function. Here are some properties of h : $(H1)$ $h(0) = 0$ and $h'(0) = 1$; $(H2)$ $h'(t) = 1 - \frac{t}{\sqrt{t^2+1}} > 0, \quad \forall t \in \mathbb{R}$, where $0 < h'(t) < 2, \forall t \in \mathbb{R}$; $(H3)$ $h''(t) = \frac{-1}{(t^2+1)^{\frac{3}{2}}} < 0, \quad \forall t \in \mathbb{R}$. Denote for $\mathcal{P}$, the Euclidean projection operator onto Ω. Our next proposed algorithm is based on Ramirez [2022] and Birgin et al. [2005].

Algorithm DHALA

Step 1. Let $x^0 \in \Omega$, $\lambda^0 \in \mathbb{R}^m_{++}$, $\tau \in I\!R_{++}$ and $0 < \alpha < 1$. Choose a positive sequence $\left\{\epsilon^k\right\}_{k \in I\!N} \subset \mathbb{R}_+$ satisfying $\lim_{k\to\infty} \epsilon^k = 0$.

Step 2. Solve the subproblem. Using x^{k-1} as initial approximation, $min \ l_H(x, \lambda^k, \tau)$, subject to $x \in \Omega$. The approximate minimizer x^k must be such that

$$\left\| \mathcal{P}\left(x^k - \nabla f(x^k) + \sum_{i=1}^{m} \lambda_i^k h'\left(\frac{\lambda_i^k g_i(x^k)}{\tau}\right) \nabla g_i(x^k)\right) - x^k \right\| \leq \epsilon^k. \tag{2.3}$$

https://proceedings.science/p/193730?lang=en

Step 3. Update the of Lagrange multipliers.

$$\lambda_i^{k+1} = \lambda_i^k h'\left(\frac{\lambda_i^k g_i(x^k)}{\tau}\right), \quad i = 1, ..., m. \tag{2.4}$$

Step 4. If the pair (x^k, λ^k) satisfies the stopping criteria: Then, stop.

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

For any $x \in \mathbb{R}^n$, we define the index sets $I_0(x) = \{i \in \{1, ..., m\} \mid g_i(x) = 0\}$, $I_+(x) = \{i \in \{1, ..., m\} \mid g_i(x) > 0\}$.

Remark 2.1 *Let $\{\lambda^k\}$ be a sequence generated by DHALA such that $\lambda_i^k > 0,\ i = 1, ..., m$ and let $\tau > 0$. Let us consider the following cases: (c1) If $i \in I_0(x)$, then we have at the k-th iteration that $g_i(x^k) = 0$, then by (2.4) and (H1), we get, $\lambda_i^{k+1} = \lambda_i^k$; (c2) If $i \in I_+(x)$, then we have at the k-th iteration that $g_i(x^k) > 0$, then $\lambda_i^{k+1} < \lambda_i^k$.*

3. Convergence

In this section, we are going present our main result.

Proposition 3.1 *Let $\left\{\lambda^k = (\lambda_1^k, ..., \lambda_m^k) \mid k = 1, 2, ...\right\} \subset \mathbb{R}^m$. If*

$$\lambda^k \in \mathbb{R}^m_{++} \quad then \quad \lambda^{k+1} \in \mathbb{R}^m_{++}.$$

Proof. See Proposition 3.2.1 of Ramirez [2022]. ∎

The following result is similar to Theorem 1 of Birgin et al. [2005]. From now on, we will assume that the stationarity condition is ensured.

Proposition 3.2 *There is a solution for the subproblem in* **Step 2**.

Proof. The subproblem in **Step 2** is a smooth minimization problem in the compact Ω. Therefore, their solution satisfy the condition (2.3). ∎

Theorem 3.1 *Let us consider problem (P). Assume that the whole sequence $\left\{x^k\right\}$ is generated by DHALA such that*

$$\lim_{k\to\infty} x^k = x^* \quad with \quad g_i(x^*) \geq 0,\ i = 1, ..., m,$$

and $\{\lambda^k\}$ is bounded. Then, (x^*, λ^*) is KKT point.

Proof. From (2.3) and (2.4), we have

$$\lim_{k\to\infty} \left\| \mathcal{P}\left(x^k - \nabla f(x^k) + \sum_{i=1}^m \lambda_i^k h'\left(\frac{\lambda_i^k g_i(x^k)}{\tau^k}\right) \nabla g_i(x^k)\right) - x^k \right\| = 0. \tag{3.5}$$

By the assumption of the statement, we have that $\lim_{k\to\infty} x^k = x^*$, where $g_i(x^*) \geq 0,\ i = 1, ..., m$. Since $\{\lambda^k\}$ is bounded and by Proposition 3.1, we have $\lim_{k\to\infty} \lambda^k = \lambda^* \geq 0,\ \forall k \in K$, where $K \subset I\!N$ is an infinite sequence of indices. Now taking limits in (3.5), so we obtain $\|\mathcal{P}\left(x^* - \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*)\right) - x^*\| = 0$. Now, we are going to ensure that the complementarity condition is verified. Since x^* is a feasible point, we have the following cases:

https://proceedings.science/p/193730?lang=en

(a) if $i \in I_0(x^*)$, then we have $g_i(x^*) = 0$. So, from (H1), (2.4) and Proposition 3.1, we have $\lim_{k\to\infty} \lambda_i^{k+1} = \lim_{k\to\infty} \lambda_i^k h'\left(\frac{\lambda_i^k g_i(x^k)}{\tau}\right) = \lim_{k\to\infty} \lambda_i^k h'(0) = \lambda_i^* \geq 0$.

(b) if $i \in I_+(x^*)$, then we have, $g_i(x^*) > 0$. So, there exists $\bar{k}$ such that for all $k \geq \bar{k}$, we have $g_i(x^k) > 0$. Then, by (c2) of the Remark 2.1, we have $\frac{\lambda_i^k g_i(x^k)}{\tau} > 0, \quad \forall k \geq \bar{k}$. Then, from (H3), (2.4) and Proposition 3.1, we have $0 < \lambda_i^k h'\left(\frac{\lambda_i^k g_i(x^k)}{\tau}\right) < \lambda_i^k h'(0), \ \forall k \geq \bar{k}$, which leads to $0 < \lambda_i^{k+1} < \lambda_i^k$. Then, by (H1), it follows that $\lim_{k\to\infty} h'\left(\frac{\lambda_i^k g_i(x^k)}{\tau}\right) = 1, \forall k \geq \bar{k}$, where $\lambda_i^* g_i(x^*) = 0$.

From the two analyses performed above, we obtain $\lambda_i^* g_i(x^*) = 0, \ i = 1, ..., m$. Thus, x^* is a KKT point. ■

4. Computational Illustration

Let us study the following problems:

Problem 4.1 *See HS19 of Hock and Schittkowski [1980]*

$$\min_{x\in\mathbb{R}^2} f(x) = (x_1 - 10)^3 + (x_2 - 20)^3$$

$$\begin{aligned} s.t. \quad g_1(x) &= (x_1 - 5)^2 + (x_2 - 5)^2 - 100 \geq 0, \\ g_2(x) &= -(x_2 - 5)^2 - (x_1 - 6)^2 + 82.81 \geq 0, \\ 13 &\leq x_1 \leq 100, \\ 0 &\leq x_2 \leq 100. \end{aligned}$$

Start: $x^0 = (20.1, 5.84)$ *(not feasible) and* $f(x^0) = -1808.858296$. *The solution is* $x^* = (14.095, 0.84296079)$ *and* $f(x^*) = -6961.81381$.

Problem 4.2 *See HS31 of Hock and Schittkowski [1980]*

$$\min_{x\in\mathbb{R}^3} f(x) = 9x_1^2 + x_2^2 + 9x_3^2$$

$$\begin{aligned} s.t. \quad g_1(x) &= x_1 x_2 - 1 \geq 0, \\ -10 &\leq x_1 \leq 10, \\ 1 &\leq x_2 \leq 10, \\ -10 &\leq x_3 \leq 1. \end{aligned}$$

Start: $x^0 = (1, 1, 1)$ *(not feasible) and* $f(x^0) = 19$. *The solution is* $x^* = (\frac{1}{\sqrt{3}}, \sqrt{3}, 0)$ *and* $f(x^*) = 6$.

4.1. Results

The algorithms were implemented in Python and executed on a Compaq running Microsoft Windows 10 Home, equipped with an Intel(R) Core(TM) i3-5005U CPU@2.00GHz. The stopping criteria used were based on feasibility and complementarity conditions, defined as follows: $\left\|\lambda^{k+1} g(x^k)\right\| \leq Tol$ and $\left\|max\{g(x^k), 0\}\right\| \leq Tol$, where Tol represents the tolerance level. In **Step 2**, we solved the subproblem using the L-BFGS-B (Zhu et al. [1997]).

https://proceedings.science/p/193730?lang=en

· Problem 4.1.
The initial conditions were: $Tol = 1E-8$, $x^0 = (20.1, 5.84)$, $\lambda^0 = (2, 2)$, $\tau = 3$.
The values obtained were: $x^* = (14.09500008,\ 0.84296094)$, $\lambda^* = (1099.152, 1231.747)$, $f(x^*) = -6961.813$ and $time(s) : 2.4375$.

· Problem 4.2.
The initial conditions were: $Tol = 1E-8$, $x^0 = (1, 1, 1)$, $\lambda^0 = (1, 1)$, $\tau = 2$.
The values obtained were: $x^* = (0.577,\ 1.732,\ 0.000)$, $\lambda^* = (3.000,\ 3.000)$, $f(x^*) = 6.000$ and $time(s) : 0.078125$.

We rewrote the same problems mentioned above, but tried to solver them considering the box constraints instead of inequality constraints. We considered the Broyden−Fletcher−Goldfarb−Shanno (BFGS) solver to solve the subproblem generated by DHALA. We observe that the algorithm solves Problem 4.2 (we consider other initial values), but it does not solve Problem 4.1. These experiments led us to think that the DHALA algorithm works more efficiently in conjunction with the L−BFGS−B solver.

Acknowledgments This work has been partially supported by FAPERJ (grant E-26/205.684/2022), COPPETEC Foundation, and CNPq (grants 302435/2019-0).

References

Birgin, E. G., Castillo, R. A., and Martínez, J. M. (2005). Numerical comparison of augmented lagrangian algorithms for nonconvex problems. *Computational Optimization and Applications*, 31(1):31–55.

Conn, A., Gould, N., and Toint, P. (1997). A globally convergent lagrangian barrier algorithm for optimization with general inequality constraints and simple bounds. *Mathematics of computation*, 66(217):261–288.

Hestenes, M. R. (1969). Multiplier and gradient methods. *Journal of optimization theory and applications*, 4(5):303–320.

Hock, W. and Schittkowski, K. (1980). Test examples for nonlinear programming codes. *Journal of optimization theory and applications*, 30:127–129.

Powell, M. J. (1969). A method for nonlinear constraints in minimization problems. *Optimization*, p. 283–298.

Ramirez, L. M. (2022). *The Hyperbolic Augmented Lagrangian Algorithm (The Carioca Algorithm)*. PhD thesis, Universidade Federal do Rio de Janeiro.

Xavier, A. E. (1992). *Penalização hiperbólica*. PhD thesis, DSc. Thesis-COPPE-UFRJ, Rio de Janeiro.

Xavier, A. E. (2001). Hyperbolic penalty: a new method for nonlinear programming with inequalities. *International Transactions in Operational Research*, 8(6):659–671.

Zhu, C., Byrd, R. H., Lu, P., and Nocedal, J. (1997). Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on mathematical software (TOMS)*, 23(4):550–560.

https://proceedings.science/p/193730?lang=en