



HAL
open science

L'Algorithme Lagrangien augmenté hyperbolique

Nelson Maculan, Vinícius Alcântara Gomes Reis de Souza, Lennin Mallma
Ramirez

► **To cite this version:**

Nelson Maculan, Vinícius Alcântara Gomes Reis de Souza, Lennin Mallma Ramirez. L'Algorithme Lagrangien augmenté hyperbolique. Doctorat. Brésil. 2024. hal-04753586

HAL Id: hal-04753586

<https://hal.science/hal-04753586v1>

Submitted on 25 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Public Domain

L'Algorithme Lagrangien augmenté hyperbolique

Nelson Maculan* Vinícius Alcântara G.R.S. † Lennin Mallma Ramirez‡

10/2024

Dans cette note, nous nous intéressons à résoudre le problème d'optimisation convexe avec des contraintes d'inégalité.

$$\min \{f(x) \mid x \in S\}, \quad (1)$$

où $S = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, \quad i = 1, \dots, m\}$, est l'ensemble convexe réalisable du problème (1) et où la fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est convexe, $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, m$, soient des fonctions convexes, nous supposons que f , g_i sont continuellement différentiables.

Les algorithmes Lagrangiens augmentés, également appelés algorithmes de multiplicateurs de Lagrange, sont importants pour résoudre ce type de problème. Ce type d'algorithme a été initialement étudié dans les travaux de [4], [8] et [9]. Actuellement, il existe une grande variété d'algorithmes basés sur différents types de fonctions Lagrangiennes augmentées, pour voir cette variété de fonctions, voir [2].

Dans cette note, nous allons considérer l'algorithme Lagrangien augmenté hyperbolique (HALA) ou également appelé algorithme du multiplicateur hyperbolique (HMA), pour résoudre le problème ci-dessus. Cet algorithme a été proposé initialement dans [10], [11] et [12]. Il est ensuite étudié plus rigoureusement dans le cas convexe, voir [6]. Dans des travaux récents, la théorie de la convergence du HALA a été étendue à des espaces mathématiques plus généraux, voir [2]; et aussi dans un autre article, la convergence a été assurée en considérant des conditions de qualification plus faibles, voir [1]. Cela nous motive à continuer à étudier ce nouvel algorithme.

Dans ce travail, nous nous intéressons principalement à résoudre ce problème, pour cela, nous considérons des problèmes convexes connus dans la littérature (voir [3] et [5]). L'objectif principal de notre travail était de montrer que nous résolvons les mêmes problèmes convexes étudiés dans [1], et que nous obtenons de meilleurs résultats de calcul par rapport à eux. Afin d'obtenir ces bons résultats, nous avons réalisé des tests numériques préalables, afin de trouver des paramètres initiaux appropriés.

1 Préliminaires

Tout au long de notre travail, nous considérerons les hypothèses suivantes :

C1. L'ensemble optimal X^* est non vide, fermé, borné et, par conséquent, compact.

C2. La qualification des contraintes de Slater est satisfaite, c'est-à-dire qu'il existe $\hat{x} \in S$ qui satisfait $g_i(\hat{x}) < 0$, pour $i = 1, \dots, m$.

En raison de l'hypothèse **C2**, il existe $\lambda^* \in \mathbb{R}^m$ tel que les conditions de Karush-Kuhn-Tucker (KKT) soient vérifiées, c.-à-d.,

*PESC/IM/COPPE, Universidade Federal do Rio de Janeiro, Brésil. email: maculan@cos.ufrj.br

†Politécnica, Universidade Federal do Rio de Janeiro, Brésil. email: vinicius.alcantara.20221@poli.ufrj.br

‡PESC/COPPE, Universidade Federal do Rio de Janeiro, Brésil. email: lenninmr@cos.ufrj.br

$$\nabla_x L(x^*, \lambda^*) = \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) = 0, \quad (2)$$

$$\lambda_i^* g_i(x^*) = 0, \quad i = 1, \dots, m, \quad (3)$$

$$g_i(x^*) \leq 0, \quad i = 1, \dots, m, \quad (4)$$

$$\lambda_i^* \geq 0, \quad i = 1, \dots, m. \quad (5)$$

Nous définissons la fonction de pénalité hyperbolique (HPF) pour ce problème comme suit (voir [10]).

$$P(y, \lambda, \tau) = \lambda y + \sqrt{(\lambda y)^2 + \frac{1}{\tau^2}}, \quad (6)$$

où $P : \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_{++} \rightarrow \mathbb{R}$. La HPF est équivalente à un lissage de la pénalité étudié par Zangwill ([13]).

1.1 Fonction Lagrangien Augmenté Hyperbolique

Nous définissons la Fonction Lagrangien Augmenté Hyperbolique (HALF) du problème (P) par $L_H : \mathbb{R}^n \times \mathbb{R}_{++}^m \times \mathbb{R}_{++} \rightarrow \mathbb{R}$,

$$L_H(x, \lambda, \tau) = f(x) + \sum_{i=1}^m \left(\lambda_i g_i(x) + \sqrt{(\lambda_i g_i(x))^2 + \frac{1}{\tau^2}} \right), \quad (7)$$

où $\tau > 0$ est le paramètre de pénalité. Maintenant nous allons présenter l'algorithme HALA pour résoudre le problème d'optimisation non linéaire (P). Cet algorithme est basé sur [6] et [7].

1.2 Algorithme

Étape 1. Soit $k := 0$ (initialisation). Prenez les valeurs initiales

$$\lambda^0 = (\lambda_1^0, \dots, \lambda_m^0) \in \mathbb{R}_{++}^m, \quad \alpha > 0 \text{ et } \tau \in \mathbb{R}_{++}.$$

Étape 2. Résoudre le problème de minimisation sans contraintes:

$$x^{k+1} = \underset{x \in \mathbb{R}^n}{\operatorname{argmin}} \left\{ f(x) + \sum_{i=1}^m \left(\lambda_i^k g_i(x) + \sqrt{(\lambda_i^k g_i(x))^2 + \frac{1}{(\tau^k)^2}} \right) \right\}.$$

Étape 3. Mise à jour des multiplicateurs de Lagrange:

$$\lambda_i^{k+1} = \lambda_i^k \left(1 + \frac{\lambda_i^k g_i(x^{k+1})}{\sqrt{(\lambda_i^k g_i(x^{k+1}))^2 + \frac{1}{(\tau^k)^2}}} \right), \quad i = 1, \dots, m. \quad (8)$$

Étape 4. (Mise à jour du paramètre τ). Si la séquence primale est réalisable, alors $\tau^{k+1} = \tau^k$. Sinon, $\tau^{k+1} = \alpha \tau^k$.

Étape 5. Si la paire (x^{k+1}, λ^{k+1}) satisfait un critère d'arrêt. FIN

Étape 6. $k := k + 1$. Revenir à l'Étape 2.

Tout au long de ce travail, nous supposons que $\lim_{k \rightarrow \infty} \tau_k = \tilde{\tau} > 0$. À l'**Étape 1**, les valeurs initiales sont fournies. HALA génère la séquence primale à l'**Étape 2** et les estimations des multiplicateurs à l'**Étape 3**. À l'**Étape 4**, le paramètre τ est mis à jour si la condition est violée, sinon il reste constant.

Considérons les hypothèses suivantes :

- C3.** Pour tout $\tau, \lambda > 0$, et pour tout $l < \infty$, l'ensemble de niveau $M = \{x \in \mathbb{R} : L_H(x, \lambda, \tau) \leq l\}$ est borné.
- C4.** Toute la séquence $\{x_k\}$ converge vers \tilde{x} , où \tilde{x} est supposé être un point réalisable, c'est-à-dire $g_i(\tilde{x}) \leq 0, i = 1, \dots, m$.

L'hypothèse **C3** est importante pour garantir qu'il existe une solution à l'**Étape 2**. L'hypothèse **C4** garantit que la séquence primale converge vers un point réalisable. Dans le résultat suivant, il a été affirmé que la séquence générée par l'algorithme converge vers un point KKT.

Theorem 1.1 *Testtest Le problème convexe (P) satisfait C1, C2, C3 et C4. Soient les séquences $\{x_k\}$ et $\{\lambda_k\}$ générées par HALA. Alors, tout point limite d'une sous-séquence $\{x_k\}$ et $\{\lambda_k\}$ est une paire solution optimale-multiplicateur de Lagrange pour le problème (P).*

Proof: Voir Theorem 3.3.1 du [6]. □

2 Expériences numériques

Dans cette section, nous présentons les résultats numériques de l'algorithme HALA pour des problèmes convexes. Les expériences ont été réalisées sur un ordinateur équipé d'un processeur Intel Core i5-1035G1, 8 Go de mémoire RAM et un système d'exploitation Debian GNU/Linux 12 via WSL 2. Le code a été écrit en langage C++ et compilé en utilisant GCC (version 12.2.0) avec l'option d'optimisation -O3 activée. Les critères d'arrêt considérés étaient la faisabilité $g(x) \leq 0$ et la convergence absolue des solutions $|x_{k+1} - x_k| < 10^{-5}$. La résolution des problèmes non contraints a été réalisée en utilisant l'algorithme Quasi-Newton L-BFGS-B. Pour tous les cas considérés, la valeur de $\alpha = 2$. Les tests ont été effectués sur les 16 problèmes convexes de la collection [5] avec des contraintes d'inégalités.

2.1 Expériences pour la détermination des paramètres initiaux

Dans cette série d'expériences, nous avons analysé la vitesse de convergence de la fonction objectif en fonction du nombre d'itérations de l'algorithme. Nous avons considéré $\lambda_{initial}, \tau_{initial} \in \{10^{-3}, 10^{-2}, 10^{-1}, 1, 5, 10, 25, 50, 75, 100, 250, 500, 750, 1000, 5000\}$.

De cette façon, 225 combinaisons de paramètres ont été considérés. Les résultats numériques sont présentés dans le Table 1.

Dans [1] les algorithmes ALGENCAN (Algorithme 3) et HALA (Algorithme 2) ont été comparés, voir Table 2, où nous trouverons aussi les bonnes performances de notre nouvelle implémentation de HALA (Algorithme 1). Les détails de cette nouvelle implémentation (Algorithme 1) sont dans la Table 1. En outre, l'**Algorithme 1** utilise l'algorithme L-BFGS-B pour la résolution des sous-problèmes générés, tandis que les **Algorithmes 2** et **3** utilisent le SPG.

Les FIGURES 1 et 2 illustrent la moyenne des itérations externes des combinaisons de paramètres initiaux qui n'ont produit aucune erreur sur aucun des problèmes. D'autres critères pour la visualisation d'une combinaison dans l'image étaient une limite de 20 itérations externes et $|f_{HALA}(x) - f_{réel}(x)| \leq 10^{-3}$.

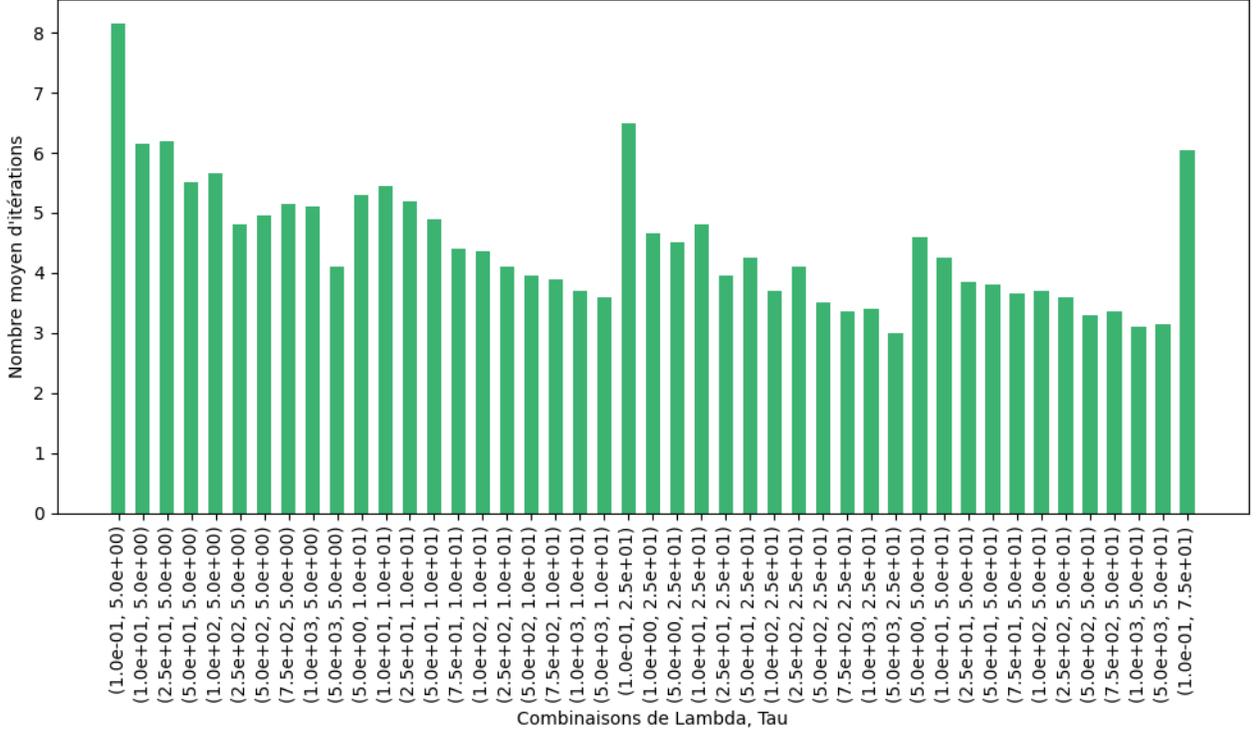


Figure 1: Moyenne des itérations par combinaison de paramètres initiaux sur les 16 problèmes (partie 1).

2.2 Expérience avec les paramètres $\lambda_{initial} = 5000$ et $\tau_{initial} = 500$

La combinaison de paramètres $\lambda_{initial} = 5000$ et $\tau_{initial} = 500$ a donné la plus petite moyenne de itérations externes sur les 16 problèmes convexes testés.

La comparaison du nombre d'itérations externes des algorithmes sur chaque problème est présentée dans le Table 2. La moyenne des itérations externes de l'**Algorithme 1** sur les 16 problèmes a été de 2.6875, tandis que les Algorithmes 2 et 3 présentaient des moyennes de 19.125 et 5.125, respectivement.

Par conséquent, ces valeurs de paramètres initiaux pour l'**Algorithme 1** garantissent moins d'itérations externes que les **Algorithmes 2** et **3**. Cela indique une amélioration significative de l'implémentation du HALA par rapport aux travaux précédents et à d'autres algorithmes lagrangiens.

References

- [1] E.G. Birgin, G. Haeser, N. Maculan, and L. Mallma Ramirez. On the global convergence of a general class of augmented lagrangian methods. 2024.
- [2] M.V. Dolgopolik. Convergence analysis of primal-dual augmented lagrangian methods and duality theory. *Preprint*, 2024.

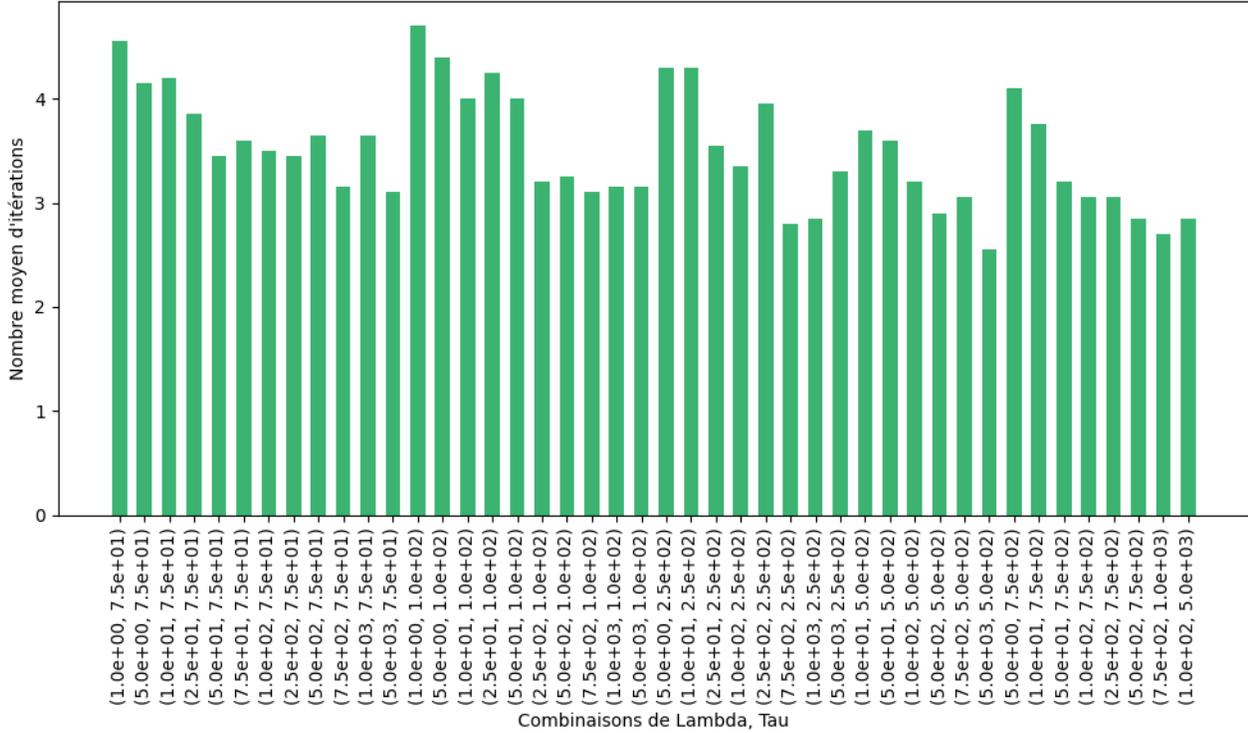


Figure 2: Moyenne des itérations des combinaisons de paramètres initiaux sur les 16 problèmes (partie 2).

Problèmes	n	m	$f(x^*)$	Itérations	Temps CPU
HS10	2	1	-1.0000001637	3	0.0030660001
HS11	2	1	-8.4984641504	3	0.0042369999
HS12	2	1	-30.0000001369	2	0.0106350007
HS21	2	1	-99.9600000000	2	0.0000800000
HS22	2	2	1.0000000128	3	0.0048070000
HS30	3	1	1.0000000000	2	0.0001459999
HS34	3	2	-0.8340323531	2	0.0348969996
HS35	3	1	0.1111107698	3	0.0025950000
HS43	4	3	-43.9999978010	2	0.0191500001
HS64	3	1	6299.8424255390	2	0.0004410000
HS65	3	1	0.9535289746	3	0.0073190001
HS66	3	2	0.5181635928	2	0.0219519995
HS72	4	2	727.6795351724	5	0.0005600001
HS76	4	3	-4.6818180712	2	0.0024510003
HS100	7	4	680.6300869724	3	0.0507009998
HS113	10	8	24.3062100519	4	0.2811560035

Table 1: Tableau avec des résultats numériques.

- [3] I. Griva and R.A. Polyak. Proximal point nonlinear rescaling method for convex optimization. *Numerical Algebra Control and Optimization*, 1:283–299, 2011.
- [4] M.R. Hestenes. Multiplier and gradient methods. *JOTA*, 4:303–320, 1969.

Problèmes	Algorithme 1	Algorithme 2	Algorithme 3
HS10	3	16	13
HS11	3	12	14
HS12	2	13	2
HS21	2	31	1
HS22	3	16	7
HS30	2	5	1
HS34	2	18	3
HS35	3	18	2
HS43	2	32	3
HS64	2	9	9
HS65	3	18	2
HS66	2	15	3
HS72	5	14	14
HS76	2	35	2
HS100	3	25	3
HS113	4	29	3

Table 2: Tableau avec de nombre des itérations.

- [5] W. Hock and K. Schittkowski. *Test Examples for Nonlinear Programming Codes*, volume 187 of *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag, Berlin, Germany, 1981.
- [6] L. Mallma Ramirez. *The hyperbolic Augmented Lagrangian algorithm*. PhD thesis, Federal University of Rio de Janeiro/COPPE, Rio de Janeiro, Brazil, 2022.
- [7] L. Mallma Ramirez, N. Maculan, A.E. Xavier, and V.L. Xavier. A non-quadratic augmented lagrangian algorithm for nonconvex programming. Technical Report ES-791/24, PESC/COPPE/UFRJ, 2024.
- [8] M.J.D. Powell. A method for nonlinear constraints in minimization problems. In R. Fletcher, editor, *Optimization*, pages 283–298. 1969.
- [9] R.T. Rockafellar. The multiplier method of hestenes and powell applied to convex programming. *JOTA*, 12:555–562, 1973.
- [10] A.E. Xavier. Penalização hiperbólica- um novo método para resolução de problemas de otimização. Master’s thesis, Federal University of Rio de Janeiro/COPPE, Rio de Janeiro, Brazil, 1982.
- [11] A.E. Xavier. *Penalização Hiperbólica*. PhD thesis, Federal University of Rio de Janeiro/COPPE, Rio de Janeiro, Brazil, 1992.
- [12] A.E. Xavier. Hyperbolic penalty: a new method for nonlinear programming with inequalities. *Intl. Trans. in Op. Res.*, 8:659–671, 2001.
- [13] W.I. Zangwill. Non-linear programming via penalty function. *Management Science*, 13:344–358, 1967.