HALA: Theory and Numerical Experiments

A. Chiacchio & L. M. Ramirez & R. S. Couto & N. Maculan

Universidade Federal do Rio de Janeiro/Escola Politécnica

alexandrebelfort@poli.ufrj.br

Politécnica UFRJ

COPPE UFRJ Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia

Abstract

A convergence result for a new Augmented Lagrangian Algorithm was recently proposed in [6], which solves convex optimization problems. This algorithm is called Hyperbolic Augmented Lagrangian Algorithm (HALA). A feature of this algorithm is that it uses a continuously differentiable function. In this work, we show the main convergence result of HALA, and we solve computationally known convex problems in the literature [4] using this new algorithm. We also perform computational experiments solving nonconvex problems. Despite having a fixed penalty parameter, HALA manages to converge to the exact solution within the precision of the computer in the computational experiments. Finally, we perform computational comparisons with another augmented Lagrangian algorithm.

Preliminaries

We are interested in studying the following optimization problem:

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

where $S = \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, i = 1, ..., m\}$, is the convex feasible set of the problem (P) and where the function $f : \mathbb{R}^n \to \mathbb{R}$ is convex, $g_i : \mathbb{R}^n \to \mathbb{R}, i = 1, ..., m$, are concave functions, we assume that f, g_i are continuously differentiable. We consider the following assumptions:

C1 The optimal set X^* is nonempty, closed, bounded and, consequently, compact.

C2 Slater constraint qualification holds, i.e., there exists $\hat{x} \in S$ which satisfies $g_i(\hat{x}) > 0, i = 1, ..., m$.

Hyperbolic Augmented Lagrangian Algorithm

The Hyperbolic Augmented Lagrangian Function (HALF) of problem (P) is defined as:

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

where $\tau > 0$ is the penalty parameter.

Algorithm

Step 1. Let $k = 0$. Take initial values $\lambda^0 = (\lambda_1^0, ..., \lambda_m^0) \in \mathrm{R}^m_{++}$ and $\tau \in \mathrm{R}_{++}$.

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

$$x^{k+1} \in argmin_{x \in \mathbb{R}^n} L_H(x, \lambda^k, \tau).$$

Step 3. Lagrange multipliers update:

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

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

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

For more details on the algorithm, see chapter 3 of [6].

Convergence Analysis

Henceforth, let us consider the following assumptions.

C3 $\forall \tau, \lambda > 0$ and $\forall l < \infty$, the level set $M = \{x \in \mathbb{R}^n \mid L_H(x, \lambda, \tau \leq l\}$ is bounded.

C4 The whole sequence $\{x^k\}$ is convergent to $\bar{x}$, where $\bar{x}$ is assumed a feasible point, i.e., $g_i(\bar{x}) \geq 0, i = 1, ..., m$.

Theorem (Theorem 3.3.1 of [6]). *The convex problem (P) satisfies **C1**, **C2**, **C3** and **C4**. Let sequences x^k and λ^k generated by HALA. Then any limit point of a sub- sequence x^k and λ^k are an optimal solution-Lagrange multiplier pair for the problem (P).*

For the full convergence analysis see chapter 3.3 of [6].

Results

Results from Table 1, compare HALA to [3] on problems from [4] denoted by HS, to [1] and [2], on versions of the Wächter-Biegler problem, to Example1 from [1] and to TP problems 1 to 5 by [5]. Problems from XIN-WEI LIU et. al, and Paul Armand and Riadh Omheni were not run on the algorithm of [3], instead, each has their algorithm.

P	n	m	AL	HALA	Found	Sol	τ	Method
HS4	2	2	0	3	2.67	2.67	1e-08	BFGS
HS5	2	4	0	2	-1.91	-1.91	1e-08	BFGS
HS10	2	1	8	7	-1.00	-1.00	1e-08	L-BFGS
HS11	2	1	9	3	-8.50	-8.50	1e-07	L-BFGS
HS12	2	1	3	7	-30.00	-30.00	1e-08	L-BFGS
HS13	2	3	26	24	0.99	1.00	1e-05	BFGS
HS19	2	6	6	2	-6961.82	-6961.81	1e-03	BFGS
HS22	2	2	8	8	1.00	1.00	1e-08	BFGS
HS25	3	6	0	3	0.01	0.00	1e-01	L-BFGS
HS30	3	7	1	8	1.00	1.00	1e-08	BFGS
HS34	3	8	3	7	-0.83	-0.83	1e-07	BFGS
HS35	3	4	3	6	0.11	0.11	1e-08	L-BFGS
HS64	3	4	10	9	6299.84	6299.84	1e-06	BFGS
HS72	4	10	14	17	727.68	727.68	1e-04	BFGS
HS76	4	7	3	9	-4.68	-4.68	1e-08	L-BFGS
HS93	6	8	6	3	135.08	135.08	1e-01	BFGS
HS108	9	14	3	2	-0.87	-0.87	1e-05	BFGS
HS113	10	8	3	17	24.31	24.31	1e-01	BFGS
WB [1]	3	6	2	2	1.00	1.00	1e-08	BFGS
WB[2]	3	6	7	2	2.00	2.00	1e-08	BFGS
Example1 [1]	2	2	14	2	-1.00	-1.00	1e-08	L-BFGS
TP1 [5]	2	2	11	7	0.99	1.00	1e-07	BFGS
TP2 [5]	2	4	10	9	-0.00	0.00	1e-08	BFGS
TP3 [5]	2	3	8	3	-0.00	0.00	1e-08	BFGS
TP4 [5]	1	2	9	2	2.00	2.00	1e-08	BFGS
TP5 [5]	2	3	16	16	0.99	1.00	1e-04	BFGS

Table 1: Results from python implementation using scipy and numpy libraries. With stopping criteria: $\|x^k - x^{k+1}\| < 10^{-3}$ and, for all problems but 43 to 45, feasibility of x^{k+1}.

On problems 43 to 45, [5] found an infeasible stationary point, so the feasibility stopping criteria of HALA had to be removed for it to stop. Even though the only criterion was the difference of consecutive x's, HALA converged to the same objective function value.

Also, since HALA compares x^k and x^{k+1}, it has a minimum of 2 iterations, and so, it is possible to see, in Table 1, that HALA stops, in several problems, within this value.

In addition, we also tested nonconvex problems: HS4, HS5, HS13, HS25, HS93, and HS108. Nevertheless, HALA managed to find the optimal solution.

Conclusion

- Despite not having a penalty parameter update step, i.e., τ is fixed, HALA has similar performance to other Augmented Lagrangian Algorithms.
- There are plenty of libraries that are made to solve the unconstrained optimization problem of **Step 2**, thus the rest of the algorithm is simple to implement.

References

[1] R. Andreani, E. G. Birgin, J. M. Martínez, and M. L. Schuverdt. On augmented lagrangian methods with general lower-level constraints. *SIAM Journal on Optimization*, 18(4):1286–1309, 2008.

[2] Paul Armand and Riadh Omheni. A mixed logarithmic barrier-augmented lagrangian method for nonlinear optimization. *Journal of Optimization Theory and Applications*, 173(2):523–547, May 2017.

[3] Leandro da Fonseca Prudente. *Inviabilidade em métodos de Lagrangiano Aumentado*. PhD thesis, Universidade Estadual de Campinas (UNICAMP). Instituto de Matemática, 2012.

[4] Willi Hock and Klaus Schittkowski. *The Test Problems*, volume 187, pages 23–127. Springer Berlin Heidelberg, Berlin, Heidelberg, 1981.

[5] XIN-WEI LIU, YU-HONG DAI, YA-KUI HUANG, and JIE SUN. A novel augmented lagrangian method of multipliers for optimization with general inequality constraints. *MATHEMATICS OF COMPUTATION*, 92(341), May 2023.

[6] Lennin Mallma Ramirez. *The Hyperbolic Augmented Lagrangian Algorithm (The Carioca Algorithm)*. PhD thesis, Universidade Federal do Rio de Janeiro, March 2022.

Acknowledgments

The second author was supported by PDR10/FAPERJ, the third by CAPES and FAPERJ and the fourth by COPPETEC Foundation.