Journal of Computational and Applied Mathematics 259 (2014) 868–876

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

An artificial fish swarm algorithm based hyperbolic augmented Lagrangian method

M. Fernanda P. Costa [a,*], Ana Maria A.C. Rocha [b], Edite M.G.P. Fernandes [b]

[a] Centre of Mathematics, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal
[b] Algorithm R & D Center, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal

ARTICLE INFO

Article history:
Received 13 February 2013
Received in revised form 22 July 2013

Keywords:
Augmented Lagrangian
Hyperbolic penalty
Artificial fish swarm
Stochastic convergence

ABSTRACT

This paper aims to present a hyperbolic augmented Lagrangian (HAL) framework with guaranteed convergence to an ϵ-global minimizer of a constrained nonlinear optimization problem. The bound constrained subproblems that emerge at each iteration k of the framework are solved by an improved artificial fish swarm algorithm. Convergence to an ϵ^k-global minimizer of the HAL function is guaranteed with probability one, where $\epsilon^k \to \epsilon$ as $k \to \infty$. Preliminary numerical experiments show that the proposed paradigm compares favorably with other penalty-type methods.

1. Introduction

We consider the problem of finding a global optimal solution of a nonconvex constrained optimization problem up to a required accuracy $\epsilon > 0$. The mathematical formulation of the problem is

$$\min_{x\in\Omega} \quad f(x) \quad \text{subject to} \quad g(x) \le 0 \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^p$ are nonlinear continuous functions and $\Omega = \{x \in \mathbb{R}^n : -\infty < l \le x \le u < \infty\}$. Problems with equality constraints, $h(x) = 0$, are reformulated into the above form using a couple of inequality constraints $h(x) - \beta \le 0$ and $-h(x) - \beta \le 0$. Since we do not assume that the functions f and g are convex, many local minima may exist in the feasible region.

Methods based on penalty functions have been used to globally solve nonconvex optimization problems [1–8]. In this type of method, the constraint violation is combined with the objective function to define a penalty function. This function aims at penalizing infeasible solutions by increasing their fitness values proportionally to their level of constraint violation. The use of a positive penalty parameter aims to balance function and constraint violation values. Tuning penalty parameter values throughout the iterative process is not an easy task. With some penalty functions, the optimal solution of the problem is attained only when the penalty parameter approaches infinity. Augmented Lagrangian (AL) penalty functions have been proposed for solving constrained global optimization problems, and their convergence properties have been derived. For most AL functions a finite penalty parameter value is sufficient to guarantee convergence to the solution of the constrained problem [9]. In [10], a global optimization method with guaranteed convergence based on the Powell–Hestenes–Rockafellar (PHR) AL function, where the exact αBB method is used to find approximate global solutions to the subproblems, is proposed. Later, the PHR function, a nonmonotone penalty parameter tuning and a gradient-based approach to solve the

* Corresponding author. Tel.: +351 253510400.
E-mail addresses: mfc@math.uminho.pt (M.F.P. Costa), arocha@dps.uminho.pt (A.M.A.C. Rocha), emgpf@dps.uminho.pt (E.M.G.P. Fernandes).

0377-0427/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cam.2013.08.017

bound constrained subproblems have been presented in [11]. The PHR function has also been combined with stochastic population-based methods, like the electromagnetism-like mechanism of optimization [12] and the artificial fish swarm (AFS) algorithm [13], to solve (1). Other proposals concerning AL functions for global optimization can be found in [14–16]. A unified theory and convergence properties of AL methods are also discussed in [17,18].

The purpose of this paper is to present an AL framework that relies on the AFS algorithm to compute a sequence of approximate global minimizers of a real-valued objective function aiming to globally solve problem (1). The AL function is a hyperbolic augmented Lagrangian (HAL) function that makes use of the well-known 2-parameter hyperbolic penalty function [19]. The convergence properties of the HAL are studied. We show that the HAL algorithm converges to an ϵ-global solution of problem (1), provided that each subproblem is globally solved up to a tolerance of ϵ^k, where $\epsilon^k \to \epsilon$, as $k \to \infty$. Further, the classical AFS algorithm [20] is improved to suit the requirements of measure theory, so that convergence to an ϵ^k-global minimizer of the HAL function with probability one is guaranteed.

The paper is organized as follows. In Section 2, the HAL paradigm and its convergence properties are presented, and Section 3 describes the improved AFS algorithm and its asymptotic convergence properties. Section 4 shows some numerical results and we conclude the paper in Section 5.

2. Hyperbolic augmented Lagrangian paradigm

The 2-parameter hyperbolic penalty function, proposed in [19], is herein used to extend its properties to a HAL function. The hyperbolic penalty is a continuously differentiable function that depends on two positive penalty parameters. This study proposes a HAL framework aiming to converge to a global solution of problem (1). The real-valued AL aims to penalize the inequality constraints while, at each iteration k of the outer cycle, an approximate global solution to the bound constrained subproblem

$$\min_{x\in\Omega} \phi^k(x) = f(x) + \sum_{i=1}^{p} \delta_i^{(k)}[g_i(x)]_+ + \tau^{(k)} \sum_{i=1}^{p} \left(g_i(x) + \sqrt{(g_i(x))^2 + (\mu^{(k)})^2} \right) \tag{2}$$

is required, for fixed values of $\delta^{(k)}$, $\tau^{(k)}$ and $\mu^{(k)}$, where $[g_i(x)]_+ = \max\{0, g_i(x)\}$, $\delta = (\delta_1, \ldots, \delta_p)^T$ is the multiplier vector associated with the constraints $g(x) \le 0$ and $\tau, \mu > 0$ are penalty parameters. These parameters have different roles: τ is the classical increasing penalty weight and μ while decreasing aims to improve the precision of the approximation.

The method used to solve the subproblem will ensure that the bound constraints are always satisfied and a global minimum is obtained. When the objective function $\phi^k(x)$ is nonconvex, a method with guaranteed convergence to a global solution is the most appropriate. The definition of an approximate global solution is used.

Definition 1 (ϵ^k-*Global Minimizer*). Let $\phi^k(x)$ be a continuous objective function defined over a bounded space $\Omega \subset \mathbb{R}^n$. The point $x^{(k)} \in \Omega$ is an ϵ^k-global minimizer of the subproblem (2) if $\phi^k(x^{(k)}) \le \min_{y\in\Omega} \phi^k(y) + \epsilon^k$, where ϵ^k is the error bound which reflects the accuracy required for the solution.

The most important issue in any AL paradigm is related with the choice of a method to compute an approximate solution to the subproblem (2). A proper choice depends on the properties of the AL function, in particular convexity and smoothness. The herein chosen method to compute an ϵ^k-global minimizer of subproblem (2), for fixed values of $\delta^{(k)}$, $\tau^{(k)}$, $\mu^{(k)}$, is a stochastic population-based algorithm, known as the AFS algorithm. The HAL algorithm for solving the problem (1) is presented in Algorithm 1.

To measure feasibility and complementarity, at iteration k, $\|V^{(k)}\|$ is used, where $\|\cdot\|$ denotes the Euclidean norm, $V_i^{(k)} = \min\{-g_i(x^{(k)}), \delta_i^{(k+1)}\}$, $i = 1, \ldots, p$, and $\delta_i^{(k+1)}$ is the multiplier that corresponds to g_i computed at $x^{(k)}$. Based on the usual paradigm [9], the first-order multiplier vector estimate is

$$\delta_i^{(k+1)} = \begin{cases} \min\left\{ \delta_i^{(k)} + \tau^{(k)} \left(1 + \dfrac{g_i(x^{(k)})}{\sqrt{\left(g_i(x^{(k)})\right)^2 + \left(\mu^{(k)}\right)^2}} \right), \delta^+ \right\}, & \text{if } g_i(x^{(k)}) > 0 \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

for $i = 1, \ldots, p$, where $\delta^+ > 0$. We note that the penalty parameter μ decreases at all iterations and parameter τ is not updated if the feasibility–complementarity measure has improved, $\|V^{(k)}\| \le \nu \|V^{(k-1)}\|$, for $\nu \in [0, 1)$. Constants $\gamma_\tau > 1$ and $\gamma_\mu < 1$ aim to increase and decrease the penalties $\tau^{(k)}$ and $\mu^{(k)}$ respectively, throughout the iterative process. We note that choosing $\gamma_\mu < 1/\gamma_\tau$, $\{\tau^{(k)}\mu^{(k)}\}$ is a bounded monotonic decreasing sequence that converges to zero.

To converge to an optimal solution of problem (1), the algorithm requires that $\{\epsilon^k\}$ defines a monotone decreasing sequence of positive values converging to ϵ as $k \to \infty$. The algorithm terminates when a feasible solution $x^{(k)}$ that satisfies the complementarity condition and has an objective function value within ϵ of the known minimum is found, i.e., when both conditions hold: $\|V(x^{(k)})\| \le 10^{-6}$ and $f(x^{(k)}) \le LB + \epsilon$, where LB denotes the smallest function value of all algorithms that found a feasible solution to (1).