Journal of Computational Mathematics
Vol.43, No.1, 2025, 143–173.
http://www.global-sci.org/jcm
doi:10.4208/jcm.2307-m2023-0016

A PERTURBED QUASI-NEWTON ALGORITHM FOR BOUND-CONSTRAINED GLOBAL OPTIMIZATION*

Raouf Ziadi[1)] and Abdelatif Bencherif-Madani

Laboratory of Fundamental and Numerical Mathematics (LMFN), Department of Mathematics, Faculty of Sciences, University Ferhat Abbas Setif 1, Setif, Algeria

Emails: ziadi.raouf@gmail.com, raouf.ziadi@univ-setif.dz, lotfi_madani@yahoo.fr

Abstract

This paper presents a stochastic modification of a limited memory BFGS method to solve bound-constrained global minimization problems with a differentiable cost function with no further smoothness. The approach is a stochastic descent method where the deterministic sequence, generated by a limited memory BFGS method, is replaced by a sequence of random variables. To enhance the performance of the proposed algorithm and make sure the perturbations lie within the feasible domain, we have developed a novel perturbation technique based on truncating a multivariate double exponential distribution to deal with bound-constrained problems; the theoretical study and the simulation of the developed truncated distribution are also presented. Theoretical results ensure that the proposed method converges almost surely to the global minimum. The performance of the algorithm is demonstrated through numerical experiments on some typical test functions as well as on some further engineering problems. The numerical comparisons with stochastic and meta-heuristic methods indicate that the suggested algorithm is promising.

Mathematics subject classification: 90C26, 90C30.

Key words: Global optimization, Limited memory BFGS method, Stochastic perturbation, Truncated multivariate double exponential distribution.

1. Introduction

In this paper we consider the following bound-constrained global optimization problem:

$$\min_{x\in D} f(x), \tag{P}$$

where D is the hyper-rectangle $\prod_{i=1}^{n} D^{(i)}, D^{(i)} = [a^{(i)}, b^{(i)}]$, $n \geq 2$ and the objective function $f(x) : \mathbb{R}^n \to \mathbb{R}$ is differentiable but not necessarily convex. The problem (P) is of interest in many real-world problems involving objective functions which are only differentiable. Numerous algorithms, depending on the regularity of f, have been already proposed, see [7,8,16,27,29]. We are concerned here with differentiable objective functions and no additional smoothness on ∇f is required. As is well known the deterministic methods [10, 16, 26, 27] guarantee theoretically their convergence to the global minimum in a finite number of iterations. However, most of them suffer from computing challenges as the problem's size is relatively high. On the other hand, the stochastic population-based algorithms [3, 9, 24] are practically the most used. Unfortunately, these methods are not based on theoretical results that guarantee their convergence to the

* Received January 20, 2023 / Revised version received April 18, 2023 / Accepted July 4, 2023 / Published online December 25, 2023 /

1) Corresponding author

global minimum and most of them are computationally expensive. Their effectiveness, due to the lack of guidance by a gradient during the searching process, is relatively inferior in terms of convergence speed for this class of problems.

In differentiable optimization, the methods that are currently in active investigation include the conjugate gradient and quasi-Newton's methods [2], which generate a sequence of points $\{x_j\}_{j\in\mathbb{N}} \subset \mathbb{R}^n$ starting from an initial point $x_0 \in \mathbb{R}^n$ following the procedure

$$x_{j+1} = x_j + t_j d_j, \tag{1.1}$$

where d_j is a descent direction for f at x_j and $t_j \in \mathbb{R}^+$ is a step-length which ensures that x_{j+1} is a feasible point with $f(x_{j+1}) \leq f(x_j)$. Without convexity, these methods are limited in applications since often only local minima are obtained. In order to escape from these local minima, several modifications of the procedure (1.1) have been proposed. Pogu *et al.* [20] have proposed, in the case where the search space is a ball, a random perturbation of the gradient method with a fixed step-length and El Mouatasim *et al.* [7,8] have proposed respectively a random perturbation of the reduced and the conditioned gradient methods for constrained global optimization where the objective function is continuously differentiable; in these works, the perturbations are governed by the standard normal distribution $\mathcal{N}(0, I_n)$. Ziadi *et al.* [30] have introduced a competitive conjugate gradient algorithm by adjusting a Gaussian perturbation strategy to a variant of the Polak-Ribière conjugate gradient method to solve bound-constrained and unconstrained global optimization problems where the function's gradient is supposed to be fully Lipschitz. However, the drawback of the aforementioned methods, especially when the dimension of the problem is relatively high, is that an important number of the generated points lie outside the feasible domain, thereby being discarded slow down the algorithm.

To tackle this problem with only a differentiable cost function, we suggest here a direct simulation of a truncated multivariate double exponential law on the hype-rectangle D. The latter actually lends itself to truncation in D more efficiently than a multivariate Gaussian law $\mathcal{N}(\mu, \sigma I_n)$. To the best of our knowledge, the use of a truncated double exponential law is new and in view of the comments in Morgan [14, pp. 100-103], it turns out to be relatively efficient in our case. The rigorous simulation procedure is carried out in Section 4 below.

Moreover, we adjust our new truncated perturbation strategy by giving a new representation of the quasi-Newton methods. We show how to use it efficiently to deal with bound-constrained global optimization problems by combining our developed perturbation strategy with a variant of L-BFGS-B (limited memory Broyden-Fletcher-Goldfarb-Shanno with boundaries) algorithm proposed by Byrd *et al.* [4]. Recall that currently the so called L-BFGS-B algorithm is one of the most efficient quasi-Newton methods for solving large-scale bound-constrained problems due to features of rapid convergence and moderate memory requirement, but is still inadequate for non-convex global optimization. Our proposed method will be called P-LBFGSB (Perturbed L-BFGS-B algorithm). Starting from a point X_0 in D, the new sequence $\{X_k\}_{k\in\mathbb{N}}$ is given by

$$X_{k+1} \in \arg\min\left\{f\big(\mathcal{G}(X_k)\big), f\big(\mathcal{P}_k^1\big), f\big(\mathcal{P}_k^2\big), \cdots, f\big(\mathcal{P}_k^r\big)\right\} \tag{1.2}$$

with

$$\mathcal{P}_k^l = \mathcal{P}_l\big(\mathcal{G}(X_k)\big), \quad l = 1, 2, \ldots, r,$$

where $\mathcal{G}(X_k)$ is the last point obtained by a few iterations using the L-BFGS-B algorithm starting from X_k and $\mathcal{P}_l(\mathcal{G}(X_k))$, for $l = 1, 2, \ldots, r$, are the stochastic perturbations of the point $\mathcal{G}(X_k)$ that are renewed independently at each iteration k, having the following truncated