Physics Letters A 378 (2014) 2831–2844

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

A perturbed martingale approach to global optimization

Saikat Sarkar [a], Debasish Roy [a,*], Ram Mohan Vasu [b]

[a] Computational Mechanics Lab, Department of Civil Engineering, Indian Institute of Science, Bangalore 560012, India
[b] Department of Instrumentation and Applied Physics, Indian Institute of Science, Bangalore 560012, India

ARTICLE INFO

Article history:
Received 19 March 2014
Received in revised form 20 June 2014
Accepted 30 July 2014
Available online 1 August 2014
Communicated by C.R. Doering

Keywords:
Local and global extremization
Martingale problem
Random perturbations
Gain-like additive updates
Chaotic dynamics

ABSTRACT

A new global stochastic search, guided mainly through derivative-free directional information computable from the sample statistical moments of the design variables within a Monte Carlo setup, is proposed. The search is aided by imparting to the directional update term additional layers of random perturbations referred to as 'coalescence' and 'scrambling'. A selection step, constituting yet another avenue for random perturbation, completes the global search. The direction-driven nature of the search is manifest in the local extremization and coalescence components, which are posed as martingale problems that yield gain-like update terms upon discretization. As anticipated and numerically demonstrated, to a limited extent, against the problem of parameter recovery given the chaotic response histories of a couple of nonlinear oscillators, the proposed method appears to offer a more rational, more accurate and faster alternative to most available evolutionary schemes, prominently the particle swarm optimization.

1. Introduction

The need for global optimization, wherein the aim is generally to determine the global extrema of possibly non-smooth, non-convex cost (or objective) functionals subject to a prescribed set of constraints, is ubiquitous across a broad range of disciplines, from dynamical systems modeling in science and engineering to drug design and delivery. The outcome of such an exercise may, inter alia, yield useful parameter information that renders the performance of a system model optimal in a sense made precise by specifying the cost functional. In the context of most practical problems, this functional could be multivariate, multimodal and even non-differentiable, which together precludes the application of a gradient-based Newton-step whilst solving the optimization problem. In contrast to finding a local extremal point, attainment of the global extremum of the cost functional, within the domain of definition (search space) of the parameters or the design variables, may be further challenged by a possibly large dimensionality of the search space. The research opportunity created by the ineffectiveness or inapplicability of the Newton search has been the fertile ground for numerous proposals for global optimization, many of which employ stochastic (i.e. random evolutionary) search with heuristic or meta-heuristic origin [11]. Some of such notable schemes include variants of the genetic algorithm (GA) [9], particle swarm optimization (PSO) [13], ant colony optimization [5] to name a few. In the task of finding the global extremum, stochastic search schemes generally score over their gradient-based counterparts [6,3], even for cases involving sufficiently smooth cost functionals wherein well defined directional derivatives can be computed. However, as long as the search is only for the nearest local extremum and the objective functional remains differentiable, gradient-based methods offer the benefit of a substantively faster convergence to the nearest extremum owing to the directional information available with the update equations. To the authors' knowledge, none of the evolutionary global optimization schemes is equipped with an equivalent of such a well defined search direction.

In organizing the global search, most evolutionary schemes depend on a random scatter applied to the available candidate solutions and the criteria to decide if the new candidates are acceptable. Such randomly drawn candidate solutions are henceforth called 'particles'. A 'greedy' selection for the new particles (i.e. choosing those offering more favorable values of the cost functional), though tempting from the perspective of faster convergence, may encounter the pitfall of getting trapped in local extrema. Most evolutionary global search methods have built-in safeguards against such premature convergence. However, despite the popular appeal of several evolutionary methods of the heuristic/meta-heuristic type [8], the justifying arguments for these schemes often draw sustenance from sociological or biological

* Corresponding author.
E-mail address: royd@civil.iisc.ernet.in (D. Roy).

http://dx.doi.org/10.1016/j.physleta.2014.07.044
0375-9601/

metaphors [11,9,13,5] that may lack a sound probabilistic basis, even though a random search always forms the cornerstone of the algorithm. Howbeit, the wide adoption of these methods owes as much to the algorithmic simplicity as to an efficient global search, which is often accomplished far more effectively than some of the more well-founded stochastic search techniques, e.g. simulated annealing [25], stochastic tunneling [26] etc. But the absence of a proper probabilistic basis for a (meta-)heuristic scheme may engender a crisis of confidence in an assuredly uniform performance across a broad range of problems and difficulties in enforcing some criterion for performance optimality. This may in turn precipitate slow convergence to the global optima and also require the end-user to tune a large set parameters for enforcing an appropriate '*exploration-exploitation trade-off*', a term used to indicate the relative balance of the random scatter (diffusion) of the particles vis-a-vis their selection based on an evaluation of the cost functional. Unfortunately, the ideal values of the tuning parameters, on which the performance of the scheme often depends crucially, may change from problem to problem. On the other hand, the concept of random evolution, based on Monte Carlo (MC) sampling of the particles, provides a means to efficient exploration of the search space, albeit at the cost of possibly slower convergence [19] in comparison with a gradient based approach. Dispensing with the notion of derivatives, as in the meta-heuristic methods, has also the added advantage of wider applicability. Many such schemes, e.g. the GA, typically have steps like 'crossover', 'mutation', 'selection' etc. While 'crossover' and 'mutation' are essentially ways of bringing in variations in the particles, the 'selection' step is used to assign to each particle a 'weight' or 'fitness' value (a measure of importance) determined as the ratio of the realized value of the cost functional, upon substitution of the particle, to the available maximum of the same. The fitness values, used to update the particles by selecting the 'best-fit' individuals for the subsequent 'crossover' and 'mutation', may be considered functionally analogous to the derivatives used in gradient based approaches. Parallel to the notion of individual fitness in the GA is that of the likelihood ratio (or weight) assigned to a particle, an MC realization at a given iteration of the state evolving as a stochastic process, in a class of nonlinear stochastic filters, e.g. the sequential importance sampling filter [10]. Although a weight based approach reduces some measure of misfit between the available best and the rest within a finite ensemble of particles, they bring with them the curse of degeneracy, i.e. all particles but one tend to assume negligibly small weights as iterations progress [1]. This problem, known as 'particle collapse in stochastic filtering, may only be arrested by exponentially increasing the ensemble size (i.e. the number of particles), a requirement that can be hardly met in practical implementations involving large dimensional states [22]. Optimization schemes that replace the weight-based multiplicative update by an additive term, which is functionally equivalent to the directional information in a Newton search, bypass particle degeneracy. PSO, one such meta-heuristic search scheme that is known to be a superior performer to most variants of the GA, utilizes an additive particle update aimed at bridging the mismatch with respect to the available best, both local and global based on the particle evolution histories, whilst still preserving the 'craziness' or scatter in the evolution. Once the desired solution is arrived at, the scatter is however expected to collapse to zero. Unfortunately, none of these methods obtain the directional information in a rigorous or optimal manner, which is perhaps responsible for a painstakingly large number of functional evaluations.

In an attempt at framing a probabilistic setting that incorporates a rigorously derived directional update of the additive type, this article demonstrates that the problem of optimization may be generically posed as a martingale problem in the sense of [23], which must however be randomly perturbed to facilitate the global search. Specifically, the local extremization of a cost functional, posed as a martingale problem, is realized through the solution of an integro-differential equation, which, in line with the Freidlin-Wentzell theory [7], is randomly perturbed so that the global extremum is attained as the perturbation vanishes asymptotically. The martingale problem, whose solution obtains a local extremum of the cost functional, involves an error function, henceforth called the innovation function, which is viewed as a stochastic process parametered through the iterations and must be driven to a zero-mean martingale. Roughly, a martingale is a stochastic Markov process whose mean in a future iteration conditioned on its current value (which is a random variable) is identical with its current conditional mean [18]. The zero-mean martingale structure of the innovation then essentially implies that small perturbations of the argument vector (the design variables) in the current iteration do not change the mean of the computed cost functional in future iterations and hence the argument vector corresponds to a local extremum. It may be noted that, even though the original extremization problem is posed in a deterministic setting, the cost functional as well as its argument vector are treated as stochastic diffusion processes. In order to realize a zero-mean martingale structure for the innovation, the particles are modified based on a change of measures effected through an additive gain-type update strategy. Thus each particle from the available population is iteratively guided by an additive correction term so as to locally extremize the given cost functional. The gain coefficient, which is a replacement for and a generalization over the Frechet derivative of a smooth cost functional, provides an efficacious directional search without requiring the functional to be differentiable. In order to accomplish the global search, we propose an annealing-type update and a random perturbation strategy (herein referred to as 'scrambling') which together efficiently ensure against possible trapping of particles in local extrema.

The rest of the paper is organized as follows. Section 2 elaborates on the local optimization posed as a martingale problem leading to an integro-differential equation whose solution is a local optimum. In Section 3, the integro-differential equation is discretized and weakly solved within the MC setting so as to get around its inherent circularity. Section 4 discusses the proposed random perturbation schemes to arrive at the global optimum without getting stuck in local traps. A pseudo-code is also included in this section for clarity of the exposition. In Section 5 we compare the performance of the proposed optimization method with particle swarm optimization (PSO) in the context of extracting parameters of chaotic oscillators based on some sparse data. Finally, the conclusions of this study are drawn in Section 5.

2. Local optimization as a martingale problem

In this section, the functional extremization is posed as a martingale problem, which also includes a generic way of satisfying a given set of constraints. However, before adopting a stochastic framework, a few general remarks on the expected functional features of the new evolutionary optimization scheme would be in order.

(i) The iterative solution is a random variable (defined on the search space) over every iteration.
(ii) Thus, along the iteration axis, the solution process is considered a stochastic process, whose mean should evolve over iterations to the optimal solution.