Probabilistic Engineering Mechanics 46 (2016) 37–47

Contents lists available at ScienceDirect

Probabilistic Engineering Mechanics

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

A stochastically evolving non-local search and solutions to inverse problems with sparse data

Mamatha Venugopal [a], Ram Mohan Vasu [a], Debasish Roy [b,*]

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

ARTICLE INFO

Article history:
Received 7 February 2016
Accepted 8 September 2016
Available online 10 September 2016

Keywords:
Global optimization
Inverse problems
Martingales
Bayesian games

ABSTRACT

Building on a martingale approach to global optimization, a powerful stochastic search scheme for the global optimum of cost functions is proposed using change of measures on the states that evolve as diffusion processes and splitting of the state-space along the lines of a Bayesian game. To begin with, the efficacy of the optimizer, when contrasted with one of the most efficient existing schemes, is assessed against a family of N_p-hard benchmark problems. Then, using both simulated and experimental data, potentialities of the new proposal are further explored in the context of an inverse problem of significance in photoacoustic imaging, wherein the superior reconstruction features of a global search vis-à-vis the commonly adopted local or quasi-local schemes are brought into relief.

1. Introduction

Inverse problems aim at the recovery of unknown parameters of a system, typically a mathematical model given perhaps by a set of differential equations, based on a few noisy measurements of the system response. Solutions to inverse problems may yield crucial parameter information with potential applications in many areas of science and engineering. Despite the exciting possibilities, a generally agreed numerical framework enabling acceptable solutions to inverse problems remains elusive, partly owing to so called non-uniqueness, as it arises in a deterministic setting (a regularized quasi-Newton method to wit) engendered by model and data (measurement) insufficiency. Moreover, presence of noise in the data may cause such solutions to drift to infeasible regions. A basic recipe for solving an inverse problem is the minimization of an objective functional that specifies the misfit between the available measurements and the predictions from the recovered model. In a deterministic setup involving sufficiently smooth fields, a common way to perform this minimization is through a gradient-based local search as exemplified, say, by the iterative Gauss-Newton (GN) method [1]. A GN-based scheme necessarily incorporates certain regularization strategies [2] that impose *a priori* constraints on the inverse problem to yield stable and meaningful solutions. Here the choice of 'right' regularization parameters adds to the computational burden brought in by Jacobian calculations in nonlinear problems. Indeed most objective functions, being non-convex, multimodal and perhaps non-differentiable, preclude the very applicability of a GN-like scheme. In contrast, a Bayesian search scheme [3] founded on the probability theory affords a more natural means to account for the numerous possible solutions by allowing the underlying probability distribution to be multimodal. Starting with an assumed prior, the aim of such a scheme would be to estimate the posterior parameter distribution conditioned on the noisy measurements. An approach that incorporates Bayesian updates is based on the filtered martingale problem [4,5] wherein the parameter to be recovered is treated as a stochastic process [6], possibly with respect to an iteration variable in case the system is time-independent. It has been shown that this approach enables obtaining additive updates to the parameters based on a change of measures so as to drive the resulting measurement-prediction misfit to a zero-mean martingale [6,7]. Convergence to a martingale structure ensures that the expectation of the measurement-prediction misfit, treated as a stochastic process, will remain zero and invariant to random perturbations during subsequent iterations or temporal recursions. It is known that, under fairly general conditions, the solution to a filtered martingale problem is unique [8] and this is perhaps a welcome departure from the non-uniqueness issues that confront a deterministic setup. Nevertheless, given that solutions could be highly sensitive to data noise, model errors and varying dimensions of the data and parameter sets, a filtered martingale problem, numerically implemented through a Monte Carlo scheme involving a finite ensemble, may at best ensure that the objective function attains an available local minimum. In general, upon averaging over a multi-modal posterior distribution irrespective of whether the different modes are physically relevant or not, the

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

http://dx.doi.org/10.1016/j.probengmech.2016.09.003
0266-8920/

recovered estimates for the system parameters could be in significant error. This problem is exacerbated with sparse data availability, a case often encountered in practice. A more rational strategy could be in the form of a stochastically founded non-local or global search scheme to pick out the most relevant mode in the posterior distribution or, perhaps to redefine a modified distribution around this mode and thus address the deterioration of the quality of solutions owing to averaging over multiple possibilities. This is an important point and is precisely what the 'coalescence' strategy, described in Section 3 of this article as part of the global search, aims to achieve.

Numerous heuristic and meta-heuristic global optimization schemes [9] abound the literature, prominently including genetic algorithm (GA) [10], simulated annealing [11], particle swarm optimization (PSO) [12], differential evolution (DE) [13] and covariance matrix adaptation evolution strategy (CMA-ES) [14] to name a few. Most such evolutionary schemes begin with a random scatter of candidate solutions, henceforth referred to as particles, which evolve over subsequent iterations according to a scheme-specific update strategy. The update steps aim at enabling the particles to explore the state space in order to detect the global minimum of the objective function. Success depends on the right *exploration-exploitation* trade-off that charts out a middle path between computationally expensive exploration and quickly identifying the global extremum from amongst the available extrema. While a full exposition of various schemes is not within our scope, brief outlines of a few prominent ones should be in order. The GA and CMA-ES assign weights to each particle based on the 'closeness' of the computed objective function to its available optimal value. In particular, only the best fit particles spawn new ones at a subsequent iteration. Such a weight-based approach that neglects the 'bad' particles might lead to a faster yet premature convergence to a local extremum despite the exploratory steps involved. This problem, known as 'particle collapse' in the stochastic filtering parlance, occurs when the entire weight is assumed by a single particle as the iterations progress. This is clearly demonstrated for the case of CMA-ES while attempting to minimize some of the benchmark objective functions in Section 5. In partial amelioration of this bottleneck, schemes like DE and PSO apply heuristically derived additive corrections to particles in the update stage.

Interestingly, none of the schemes discussed so far are naturally equipped to handle multivariate/multi-objective optimization (MOO), the *sine qua non* in solving many inverse problems. Although there have been attempts at adapting the CMA-ES, a powerful evolutionary scheme, for MOO, only limited success in applications has accrued [15]. In [16,17], we have proposed a generalized optimization framework, COMBEO (Change Of Measure Based Evolutionary Optimization), based mainly on a perturbed martingale problem that could rationally accommodate updates by different existing methods within a single mathematical structure. The need for such a unified framework was inspired by the no free lunch theorems [18] that proved the near impossibility of a single optimization scheme performing well across the spectrum of N_p-hard problems. Yet another advantage of COMBEO was in its inherent ability to treat multi-objective problems. On the downside, COMBEO either required a large ensemble size or an inflated number of measurements to solve a given problem. One of our current aims is thus to modify COMBEO so as to better equip it to solve practical problems with smaller ensemble sizes and sparser sets of measurements. This is primarily accomplished by incorporating within the martingale problem of local optimization, a new update strategy based on state space splitting (3S). Additionally, perturbative exploratory steps such as scrambling, blending etc. guide the greedy local search to converge to the global optimum. Our second focus here is on establishing, using experimental data, the efficacy of a non-local search as encoded within COMBEO for parameter reconstruction and contrast such performance with that of a more localized search as represented, say, by COMBEO stripped off its global search tools. In the process, we demonstrate the usefulness of the 3S scheme in solving inverse problems with sparse measurements.

The rest of the paper is organized as follows. Section 2 poses optimization as a filtered martingale problem and puts forth the bare-bones additive update strategy that renders the measurement-prediction misfit a zero-mean martingale. The random exploratory operations that aid in the global search are briefly explained in Section 3. Much of the material presented in Sections 2 and 3 (with the exceptions of the blending strategy and a few additional insights) is reported elsewhere [16,17] and is included here for completeness. Section 4 contains a game theoretic perspective that leads to the 3S scheme followed by a pseudo-code for the proposed evolutionary search. The first part of Section 5 gives comparative results of the proposed scheme vis-à-vis CMA-ES in minimizing a few benchmark objective functionals. Thereafter, we undertake a numerical study to contrast the present approach vis-à-vis one without the trappings of global search in the context of a medical imaging problem, viz. quantitative photoacoustic tomography. The aim is to bring out, perhaps for the first time, what a well-conceived global search scheme can do when the available measurements are very sparse. Reconstructions for both simulated and experimental data are given. The concluding remarks are presented in Section 6.

2. Optimization as a filtered martingale problem

For a nonlinear objective functional $f(\mathbf{x}):\mathbb{R}^{n_x} \to \mathbb{R}$, our aim is to find $\mathbf{x}^{\min} \in \mathbb{R}^{n_x}$ such that $f\left(\mathbf{x}^{\min}\right) \le f(\mathbf{x})\ \forall\ \mathbf{x} \in \mathbb{R}^{n_x}$. For the multi-objective case with $\mathbf{f}(\mathbf{x}) := \left(f^1(\mathbf{x}), \dots, f^{n_f}(\mathbf{x})\right) \in \mathbb{R}^{n_f}$, each component $f^i(\mathbf{x})\ \forall\ i = 1, \dots, n_f$ has to be minimized. If we wish to solve the deterministically posed optimization problem by borrowing ideas from stochastic filtering, the parameters to be recovered and the objective functions must be treated as stochastic processes, which are possibly of the diffusion type evolving with respect to a time-like iteration variable τ. Within a complete probability space $\left(\Omega, \mathcal{F}, \mathcal{F}_\tau, P\right)$ [6], such a characterization would render $\mathbf{x}: \Omega \to \mathbb{R}^{n_x}$ an n_x-dimensional random vector at each iteration with $\mathcal{F}$ denoting the Borel σ-algebra over open subsets of $\mathbb{R}^{n_x}$, $\mathcal{F}_\tau$ the natural filtration and P the probability measure. Noting however that the parameters are usually not governed by any dynamics, $\mathbf{x}$ may have to be artificially evolved as a continuous Brownian motion or discrete random walk in τ (evolution with jump discontinuities, modeled as Levy processes, is possible and even desirable for more efficient non-local search; but not considered here). The evolution of the continuously parameterized stochastic process $\mathbf{x}_\tau$ is then represented in the form of a stochastic differential equation (SDE) [6]:

$$d\mathbf{x}_\tau = d\mathbf{B}_\tau \tag{2.1}$$

Here, $\mathbf{B}_\tau \in \mathbb{R}^{n_x}$ is a vector Brownian motion with mean zero and covariance matrix $\Sigma_B \Sigma_B^T \in \mathbb{R}^{n_x \times n_x}$. Although τ is by definition monotonically increasing in $\mathbb{R}^+$, in practice, $\mathbf{x}_\tau$ is evolved only over finite increments of τ, i.e. for $\tau_1 < \dots < \tau_{n_t}$. Thus, for the $(k+1)^{th}$ iteration where $\tau \in (\tau_k, \tau_{k+1}]$, we can write a discrete form of Eq. (2.1) as

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta\mathbf{B}_k \tag{2.2}$$

where $\Delta\mathbf{B}_k = \mathbf{B}_{k+1} - \mathbf{B}_k$. Let a minimum of the objective function be denoted as $\tilde{\mathbf{f}} = \mathbf{f}(\mathbf{x}^{\min})$. In a deterministic setup, when the design