Journal of Computational Science 16 (2016) 140–155

Contents lists available at ScienceDirect

Journal of Computational Science

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

Multi-objective constrained black-box optimization using radial basis function surrogates

Rommel G. Regis

Department of Mathematics, Saint Joseph's University, Philadelphia, PA 19131, United States

ARTICLE INFO

Article history:
Received 30 December 2015
Received in revised form 29 May 2016
Accepted 30 May 2016
Available online 6 June 2016

Keywords:
Multi-objective optimization
Constrained optimization
Expensive black-box optimization
Surrogate model
Radial basis function

ABSTRACT

This article presents a framework for a surrogate-based stochastic search algorithm for multi-objective and constrained black-box optimization where the objective and constraint function values are outputs of computationally expensive computer simulations. Unlike many other approaches, the proposed framework is *not* population-based and handles constraints without explicitly using a penalty function. In each iteration, the algorithm constructs or updates response surface models or surrogate models of the objective and constraint functions. Then, it generates multiple random trial points according to some probability distribution over the search space. The surrogate models for the objective and constraint functions are then used to identify the trial points that are predicted to be feasible and nondominated. From this set of trial points, two criteria are used to select the next sample point where the expensive objective and constraint functions will be evaluated. These criteria are the minimum distance of the predicted objective vector of a trial point from the current set of nondominated objective vectors and also the minimum distance of the trial point from previous sample points. The proposed framework is implemented using radial basis function (RBF) surrogate models and compared with alternative methods, including NSGA-II and Uniform Random Search on 28 benchmark test problems. The numerical results indicate that the proposed method is promising for computationally expensive multi-objective and constrained black-box optimization.

1. Introduction

This article develops a framework for a surrogate-based stochastic algorithm for multi-objective constrained optimization that can be used for problems with computationally expensive black-box objective and constraint functions. In these problems, the values of the objective and constraint functions for given settings of the input variables are obtained via time-consuming computer simulations that could take many hours per simulation. Consequently, the total time spent on simulations completely dominates the running time of a typical optimization algorithm on these computationally expensive problems. These problems are found in many engineering applications, particularly those involving finite element or computational fluid dynamics simulations (e.g., Prieß et al. [1], Clees et al. [2], Bureerat and Srisomporn [3], Husain et al. [4]).

Specifically, this article proposes a class of algorithms for the following multi-objective constrained optimization problem:

$$\begin{aligned} &\min \quad F(x) = (f_1(x), \ldots, f_k(x)) \\ &\text{s.t.} \\ &\quad G(x) = (g_1(x), \ldots, g_m(x)) \le 0 \\ &\qquad\qquad \ell \le x \le u \end{aligned} \tag{1}$$

Here, $\ell, u \in \mathbb{R}^d$ and the functions $f_i : \mathbb{R}^d \to \mathbb{R}$, $i = 1, \ldots, k$ and $g_j : \mathbb{R}^d \longrightarrow \mathbb{R}$, $j = 1, \ldots, m$ are black-box in that their mathematical forms are not explicitly available and instead their values are obtained via an expensive but deterministic simulation. Also, one *simulation* for a given input vector $x \in [\ell, u]$ yields the values of all the components of $F(x)$ and $G(x)$. The region $[\ell, u]$ defined by the bounds is referred to as the *search space* for problem (1). Moreover, the derivatives of the objective and constraint functions are assumed to be unavailable. For simplicity, assume also that there are no equality constraints and that the feasible region $\mathcal{D} := \{x \in \mathbb{R}^d : \ \ell \le x \le u, \ \ G(x) \le 0\}$ has a nonempty interior. Furthermore, assume that a feasible starting point is given, which

E-mail address: rregis@sju.edu

http://dx.doi.org/10.1016/j.jocs.2016.05.013
1877-7503/

is not unreasonable since it is often the case that an engineer knows of a feasible design to the problem and is simply looking for an improved solution. For convenience, this problem will be referred to as *MCOP*(F, G, $[\ell, u]$). Future work will deal with more difficult cases not covered here, including optimization problems with black-box equality constraints, the absence of feasible starting points, and the presence of noise in the objective and/or constraint functions.

The proposed framework for a stochastic algorithm for problem (1) dynamically updates response surface models or surrogate models for each of the objective and constraint functions at the beginning of each iteration. Then, it randomly generates multiple trial points, called *candidate points*, according to some probability distribution (e.g., uniform random over the search space or Gaussian centered at a current nondominated sample point). The algorithm then uses the surrogate models for the constraint functions to predict which candidate points will be feasible or have the minimum number of constraint violations. From this set of candidate points, the most promising point (or points if run in parallel) is identified according to multiple criteria such as predicted non-domination rank according to the surrogates, distance from previously evaluated points in the decision space, and distance of the predicted objective vector from the current nondominated objective vectors in the objective space. The expensive simulation is then carried out only on this promising trial point. Unlike other methods for constrained multi-objective optimization, the proposed approach does *not* lump the constraints into a single penalty function since this is not expected to be effective in the computationally expensive setting. Moreover, it is *not* population-based but it can be easily modified to generate multiple sample points for parallel processing. The proposed framework, called *MOCS-RS (Multi-Objective Constrained Stochastic optimization using Response Surfaces)*, is an extension to the multi-objective setting of the ConstrLMSRS approach (Regis [5]) for constrained black-box optimization, which was shown to work well on a large-scale benchmark problem with 124 decision variables and 68 black-box inequality constraints.

The proposed approach is implemented using Radial Basis Function (RBF) surrogates and compared with alternative methods, including NSGA-II (Deb et al. [6]), Direct Multi Search (DMS) (Custódio et al. [7]) and a random search algorithm for multi-objective constrained optimization, on 28 benchmark test problems. The numerical experiments show that the resulting method called *MOCS-RBF* is very promising for multi-objective constrained black-box optimization.

At present, relatively few surrogate-based algorithms have been proposed for problems with multiple black-box objective functions and inequality constraint functions. Hence, one important contribution of the proposed framework is another approach for solving these problems. Moreover, the approach in this paper differs from that of others in that a framework is being proposed instead of specific algorithms. This framework is meant to encompass various types of surrogate models (including ensembles), probability distributions that generate the candidate points, and criteria or strategies for selecting the sample point from the set of candidate points. This provides users with the freedom to explore schemes within the framework that they think might be effective. Another valuable contribution of this paper is an extensive numerical test to compare implementations of MOCS-RBF with alternative methods. A majority of papers on surrogate-based optimization test the proposed methods on only a handful of generally low-dimensional problems while this paper uses 28 test problems with up to 15 decision variables, up to 5 objectives and up to 11 inequality constraints. Also, each algorithm being tested is run 30 times on each problem. In addition to simply using the widely popular hypervolume metric, this paper uses data profiles [8] to compare how well the nondominated objective vectors are spread out across the approximate Pareto front. Hence, this paper presents more reliable numerical results that confirm that robustness of the specific algorithms that this article proposes.

2. Preliminaries and notations

Before proceeding, a few terms need to be defined in the context of multi-objective constrained optimization. Consider an MCOP(F, G, $[\ell, u]$) of the form (1). This article employs some terminology from standard texts in multi-objective optimization (e.g., Miettinen [9]). Below are some basic terms. In the definitions below, $\mathcal{D}$ is the feasible region of problem (1).

Definition 2.1. A point $x \in \mathcal{D}$ dominates another point $y \in \mathcal{D}$, written $x \prec y$, if $f_i(x) \le f_i(y)$ for all $i = 1, \ldots, k$ and $f_j(x) < f_j(y)$ for some j.

Definition 2.2. A point $x^* \in \mathcal{D}$ is a (global) *Pareto minimizer* of F over $\mathcal{D}$ if $\nexists y \in \mathcal{D}$ s.t. $y \prec x^*$. The *Pareto set of F over* $\mathcal{D}$, denoted by $\mathcal{X}^*_{F,\mathcal{D}}$, is the set of all global Pareto minimizers of F over $\mathcal{D}$. The *Pareto front of F over* $\mathcal{D}$ is the image of the Pareto set under the mapping F, i.e., it is $F(\mathcal{X}^*_{F,\mathcal{D}}) = \{F(x^*) \ : \ x^* \in \mathcal{X}^*_{F,\mathcal{D}}\}$.

Ideally, one would wish to determine the entire Pareto set and Pareto front of F over $\mathcal{D}$. However, for many practical problems, the Pareto set and Pareto front are infinite sets, and so, one can only hope to find a finite representative subset of these sets. In practice, many algorithms strive to find a nondominated subset of objective vectors, sometimes with no guarantee of obtaining any Pareto optimal solutions. The solutions found can then be presented to a decision maker who might select one or a few nondominated solutions for implementation.

Since this paper focuses on constrained multi-objective optimization problems, it is necessary to extend the concept of non-domination to the entire search space $[\ell, u]$, including infeasible points that satisfy the bounds. To do this, we first define the concept of a constraint violation function.

Definition 2.3. Let $[\ell, u] \subseteq \mathbb{R}^d$ be the search space and let $G(x)$ be the constraint function for problem (1). Moreover, let $\mathcal{D} = \{x \in \mathbb{R}^d : \ \ell \le x \le u, \ \ G(x) \le 0\}$ be the feasible region of the problem. A *constraint violation function* for G over $[\ell, u]$ is a function $V : [\ell, u] \to \mathbb{R}^+$ satisfying the following conditions:

(i) $V(x) = 0$ for all $x \in \mathcal{D}$;
(ii) $V(x) > 0$ for all $x \notin \mathcal{D}$; and
(iii) If $G(x) \le G(y)$, then $V(x) \le V(y)$.

A constraint violation function is a measure of the degree of constraint violation of a point in the search space $[\ell, u]$. It is easy to verify that $V(x) = \sum_{j=1}^{m} [\max\{g_j(x), 0\}]^q$, where $q > 0$, and $V(x) = \sum_{j=1}^{m} I([g_j(x) > 0])$, where $I(\cdot)$ is the indicator function, are constraint violation functions for G over $[\ell, u]$. Commonly used examples are $V(x) = \sum_{j=1}^{m} [\max\{g_j(x), 0\}]$ and $V(x) = \sum_{j=1}^{m} [\max\{g_j(x), 0\}]^2$. In fact, the NSGA-II code that is used in the comparisons uses the former constraint violation function.

Next, the concept of *domination* is extended to all points in the search space $[\ell, u]$.

Definition 2.4. Consider the MCOP(F, G, $[\ell, u]$) in (1) with feasible region $\mathcal{D}$ and let $V(x)$ be a constraint violation function for G over $[\ell, u]$. A point $x \in [\ell, u]$ dominates another point $y \in [\ell, u]$, written $x \prec y$, if any one of the following conditions hold:

(a) $x, y \in \mathcal{D}$ and $f_i(x) \le f_i(y)$ for all $i = 1, \ldots, k$ and $f_j(x) < f_j(y)$ for some j;