Computers and Chemical Engineering 93 (2016) 353–360

Contents lists available at ScienceDirect

Computers and Chemical Engineering

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

Constraint back-offs for safe, sufficient excitation: A general theory with application to experimental optimization

Gene A. Bunin

Ronin Institute for Independent Scholarship, United States

ARTICLE INFO

Article history:
Received 3 January 2016
Received in revised form 2 May 2016
Accepted 3 July 2016
Available online 5 July 2016

Keywords:
Design measures for robustness
Evolutionary operation
Optimization under uncertainties
Experiment design

ABSTRACT

In many experimental settings, one is tasked with obtaining information about certain relationships by applying perturbations to a set of independent variables and noting the changes in the set of dependent ones. While traditional design-of-experiments methods are often well-suited for this, the task becomes significantly more difficult in the presence of constraints, which may make it impossible to sufficiently excite the experimental system without incurring constraint violations. The key contribution of this paper consists in deriving constraint back-off sizes sufficient to guarantee that one can always perturb in a ball of radius δ_e without leaving the constrained space, with δ_e set by the user. Additionally, this result is exploited in the context of experimental optimization to propose a constrained version of G. E. P. Box's evolutionary operation technique. The proposed algorithm is applied to three case studies and is shown to consistently converge to the neighborhood of the optimum without violating constraints.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

In most branches of science, one often encounters systems where the relationship between some experimental response and a finite number of independent variables needs to be studied (Myers et al., 2009; Montgomery, 2012), and it is generally assumed that the response is a dependent variable and a function of the independent ones. Mathematically, the experimental quantity may be stated as the function $f : \mathbb{R}^{n_u} \to \mathbb{R}$, while $\mathbf{u} \in \mathbb{R}^{n_u}$ may be used to denote the vector of independent variables $\mathbf{u} = (u_1, \ldots, u_{n_u})$. One is then left with the task of identifying $f(\mathbf{u})$. The identification may be either local or global, and is usually done by conducting a series of experiments with different values of $\mathbf{u}$, observing the resulting $f(\mathbf{u})$ values, and performing some sort of regression. Such procedures are typically used to:

(i) construct data-driven model approximations of f for when f is difficult to model via first principles (Jones et al., 1998; Myers et al., 2009; Montgomery, 2012),
(ii) estimate the uncertain parameters of an already available model (Box, 1990; Chen and Joseph, 1987; Pfaff et al., 2006; Quelhas et al., 2013),
(iii) explore how the function value changes so as to find conditions for which the value is minimized, maximized, or equal to a certain quantity (Robbins and Monro, 1951; Box and Wilson, 1951; Lewis et al., 2000; Conn et al., 2009).

It is the case for many problems that the experimental space of interest is a box defined by the constraints $u_i^L \le u_i \le u_i^U, \quad i = 1, \ldots, n_u$, where $\mathbf{u}^L = (u_1^L, \ldots, u_{n_u}^L)$ and $\mathbf{u}^U = (u_1^U, \ldots, u_{n_u}^U)$ are the lower and upper limits on the independent variables, respectively. Such problems typically correspond to simple set-ups that do not possess major safety limitations, and where testing any variable combination in the experimental space is permissible. Obtaining knowledge about f is not difficult in such conditions, and the traditional design-of-experiments techniques (Montgomery, 2012) are perfectly appropriate here.

However, there still exists a fair share of problems – many of them corresponding to continuous or batch chemical processes (Bunin, 2016) – where additional constraints enter to reduce the experimental space in a nontrivial manner. These constraints may be expressed as the n_g inequalities

$$g_j(\mathbf{u}) \le 0, \quad j = 1, \ldots, n_g.$$

In some problems, the functions g_j may represent *experimental* relationships that, like f, can only be divined empirically. It often happens that such experimental constraints are safety or economic limitations – they could, for example, represent an upper limit on the temperature in a continuous reactor, or a lower limit on the purity of a batch-produced chemical. Despite the violation of such constraints being highly undesirable, or even dangerous,

E-mail address: gene.bunin@roninstitute.org

http://dx.doi.org/10.1016/j.compchemeng.2016.07.006
0098-1354/© 2016 Elsevier Ltd. All rights reserved.

there currently exists no easy-to-implement, theoretically rigorous method for guaranteeing that the perturbations carried out on the system satisfy these constraints.

Notably, there does exist a fairly established literature on methods that suppose the existence of a parametric model approximation $g_{m,j}(\mathbf{u},\boldsymbol{\theta}) \approx g_j(\mathbf{u})$, define Θ as the uncertainty set to which $\boldsymbol{\theta}$ belongs, and then attack the problem via probabilistic formulations by ensuring that $g_{m,j}(\mathbf{u},\boldsymbol{\theta}) \leq 0$ with sufficiently high probability (Kall and Wallace, 1994; Zhang et al., 2002; Sahinidis, 2004; Li et al., 2008; Quelhas et al., 2013). However, while such methods are theoretically just and robust, they suffer from four major practical drawbacks:

(i) the requirement of a parametric model,
(ii) the restriction that the uncertainty be parametric, and that Θ be known,
(iii) the computational issues that arise with probabilistic constraints,
(iv) the conservatism that results from the probabilistic constraints reducing the set of admissible $\mathbf{u}$.

Drawback (i) becomes debilitating when the system at hand is difficult to model, while (ii) is more problematic since many employed models are, often by practical requirement, simplifications and thereby prone to *structural* errors (Chachuat et al., 2009). Drawback (iii) is likely to be significant when the models have many decision variables, many uncertain parameters, and are involved. Simplifications, such as linearizing the model with respect to $\boldsymbol{\theta}$ (Zhang et al., 2002), may be used to avoid this, but ultimately come with the loss of rigor that one would expect from an approximation. Finally, (iv) can be extremely problematic when the parametric uncertainty set is large – as may often occur in practice (Li et al., 2008; Quelhas et al., 2013) – since this may limit the perturbation options, with only a small collection of $\mathbf{u}$ being deemed "safe".

The methodology proposed in the present work avoids these difficulties while maintaining the rigor. Taking a model-free, back-off approach, we simplify and generalize the results of Bunin et al. (2014a) to derive positive values, b_j, that, for a given $\mathbf{u}^*$, allow us to state the guarantee

$$g_j(\mathbf{u}^*) \leq -b_j \Rightarrow g_j(\mathbf{u}) \leq 0, \quad \forall \mathbf{u} \in \mathcal{B}_e, \tag{1}$$

where

$$\mathcal{B}_e = \{\mathbf{u} : \|\mathbf{u} - \mathbf{u}^*\|_2 \leq \delta_e\}.$$

Verbally, this means that given a decision-variable set $\mathbf{u}^*$ known to satisfy the constraints with some slack, one is able to provide a guarantee that the entire ball of radius δ_e surrounding $\mathbf{u}^*$ will satisfy the constraints as well, thereby allowing the user to perturb anywhere within this ball without fear of constraint violation. Despite being local, such a result is nevertheless very useful as it allows a high degree of freedom – a ball permitting perturbation sets of any geometry. As will be shown, the value b_j will depend on the local sensitivities of g_j around $\mathbf{u}^*$, but can nevertheless be computed without requiring much effort from the user. Conversely, δ_e is the *sole tuning parameter* set by the user and represents, in some sense, the magnitude of perturbation considered as "sufficiently exciting" for identification given the particular problem.

To date, this result has already been integrated into the SCFO experimental optimization solver (Bunin, 2015), where it is used to ensure accurate linear and quadratic regression, but it is expected that the generality of the result make it applicable to many algorithms and contexts. In this paper, its usefulness is illustrated for a much simpler optimization algorithm – the evolutionary operation (EVOP) method of Box (Box, 1957; Box and Draper, 1969). As the original method searches to maximize an experimental function by perturbing in a hypercube around the best known $\mathbf{u}$, it is made coherent with the result here by ensuring that the cube lie inside $\mathcal{B}_e$, with $\mathbf{u}^*$ then defined as the best known reference point. By forcing $\mathbf{u}^*$ to always satisfy (1), it thus follows that all exploration by the modified EVOP version satisfy the constraints.

The remainder of this paper is organized as follows. The required mathematical concepts and the derivation of the appropriate constraint back-offs are presented in Section 2. Section 3 then provides a robust extension of (1) that accounts for noise/error in the function values, together with a general discussion of potential implementation issues. The constrained EVOP algorithm is presented in Section 4, and its effectiveness is illustrated for three case-study problems. Section 5 concludes the paper.

2. Derivation of sufficient back-offs

So as to keep the forthcoming analysis relatively simple, the following assumption on the continuity and differentiability of g_j is made.

Assumption 1. The functions g_j are continuously differentiable ($\mathcal{C}^1$) on an open set containing $\mathcal{B}_e$.

This then allows for the definition of bounds on the sensitivities of g_j.

Definition 1. The local Lipschitz constants of g_j are defined as any constants κ_{ji} satisfying

$$-\kappa_{ji} \leq \left.\frac{\partial g_j}{\partial u_i}\right|_{\mathbf{u}} \leq \kappa_{ji}, \quad \forall \mathbf{u} \in \mathcal{B}_e. \tag{2}$$

The existence of these constants follows from Assumption 1 and the boundedness of $\mathcal{B}_e$. They may be used to bound the violation of a given g_j via the local *Lipschitz upper bound*.

Lemma 1. *Let $\mathbf{u}_a, \mathbf{u}_b \in \mathcal{B}_e$. It follows that*

$$g_j(\mathbf{u}_b) \leq g_j(\mathbf{u}_a) + \sum_{i=1}^{n_u} \kappa_{ji} |u_{b,i} - u_{a,i}|. \tag{3}$$

Proof. See Bunin et al. (2014b). □

Finally, the Lipschitz bound may be exploited by substituting $\mathbf{u}_a \to \mathbf{u}^*$ and $\mathbf{u}_b \to \mathbf{u}$ in (3) to generate a *Lipschitz polytope* around the point $\mathbf{u}^*$.

Definition 2. Let $\mathcal{L}_j$ denote the Lipschitz polytope of the constraint g_j around $\mathbf{u}^*$, defined as the set

$$\mathcal{L}_j = \left\{ \mathbf{u} : g_j(\mathbf{u}^*) + \sum_{i=1}^{n_u} \kappa_{ji} |u_i - u_i^*| \leq 0 \right\}. \tag{4}$$

The Lipschitz polytope has two important properties that should be apparent by inspection:

(i) $\mathbf{u} \in \mathcal{L}_j \cap \mathcal{B}_e \Rightarrow g_j(\mathbf{u}) \leq 0$,
(ii) $g_j(\mathbf{u}^*) \leq 0 \Rightarrow \mathcal{L}_j \neq \emptyset$.

The "double membership" of $\mathbf{u} \in \mathcal{L}_j$ and $\mathbf{u} \in \mathcal{B}_e$ in (i) is required to ensure both that $\mathbf{u}$ satisfies the upper bound of (3) and that the bound itself is valid to begin with, respectively. Property (ii) should be evident if one just considers $\mathbf{u} := \mathbf{u}^*$ when $g_j(\mathbf{u}^*) \leq 0$.

Furthermore, it is clear that the content (hypervolume) of $\mathcal{L}_j$ increases monotonically as $g_j(\mathbf{u}^*)$ decreases – i.e., $\mathcal{L}_j$ admits more and more implementable points because of the terms $|u_i - u_i^*|$ being allowed to grow larger while satisfying the inequality.

It is this observation that inspires the foundations of the present work, illustrated geometrically in Fig. 1. If $g_j(\mathbf{u}^*)$ can be forced to

Download English Version:

https://daneshyari.com/en/article/172002

Download Persian Version:

https://daneshyari.com/article/172002

Daneshyari.com