Automatica 63 (2016) 198–208

Contents lists available at ScienceDirect

Automatica

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

Generalized derivatives of dynamic systems with a linear program embedded☆

Kai Höffner, Kamil A. Khan, Paul I. Barton

Process Systems Engineering Laboratory, Massachusetts Institute of Technology, Cambridge MA 02139, USA

ARTICLE INFO

Article history:
Received 3 July 2014
Received in revised form
29 June 2015
Accepted 9 September 2015

Keywords:
Generalized derivatives
Nonsmooth dynamic systems
Nonsmooth optimization
Lexicographic derivative
Linear programming

ABSTRACT

Dynamic systems with a linear program (LP) embedded can be found in control and optimization of bioreactor models based on dynamic flux balance analysis (DFBA). Derivatives of the dynamic states with respect to a parameter vector are essential for open and closed-loop dynamic optimization and parameter estimation of such systems. These derivatives, given by a forward sensitivity system, may not exist because the optimal value of a linear program as a function of the right-hand side of the constraints is not continuously differentiable. Therefore, nonsmooth analysis must be applied which provides optimality conditions in terms of subgradients, for convex functions, or Clarke's generalized gradient, for nonconvex functions. This work presents an approach to compute the necessary information for nonsmooth optimization, *i.e.*, an element of the generalized gradient. Moreover, a numerical implementation of the results is introduced. The approach is illustrated through a large-scale dynamic flux balance analysis example.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Consider a parameter-dependent system of ordinary differential equations (ODE) of the form

$$\begin{aligned}\dot{\mathbf{x}}(t,\mathbf{p}) &= \mathbf{f}(t,\mathbf{x}(t,\mathbf{p}),\mathbf{p},h(\mathbf{x}(t,\mathbf{p}),\mathbf{p})), \quad \forall t\in(t_0,t_f], \\ \mathbf{x}(t_0,\mathbf{p}) &= \mathbf{f}_0(\mathbf{p}),\end{aligned} \tag{1}$$

where $h(\mathbf{x}(t,\mathbf{p}),\mathbf{p})$ is the optimal value of the state- and parameter-dependent linear program

$$\begin{aligned}h(\mathbf{x}(t,\mathbf{p}),\mathbf{p}) = \min_{\mathbf{v}\in\mathbb{R}^{n_v}} \ & \mathbf{c}^{\mathrm{T}}\mathbf{v} \\ \text{s.t.} \ \ \mathbf{A}\mathbf{v} &= \mathbf{b}(\mathbf{x}(t,\mathbf{p}),\mathbf{p}), \\ \mathbf{v} &\geq \mathbf{0}.\end{aligned} \tag{2}$$

We denote Eq. (1) with h given in Eq. (2) as a *dynamic system with a linear program embedded*. A broad class of dynamic optimization problems for this type of system can be formulated as

$$\min_{\mathbf{p}} J(\mathbf{p}) = \int_{t_0}^{t_f} w(\mathbf{x}(\tau,\mathbf{p}),\mathbf{p})d\tau \tag{3}$$

$$\text{s.t. } \dot{\mathbf{x}}(t,\mathbf{p}) = \mathbf{f}(t,\mathbf{x}(t,\mathbf{p}),\mathbf{p},h(\mathbf{x}(t,\mathbf{p}),\mathbf{p})), \forall t\in(t_0,t_f], \tag{4}$$

$$\mathbf{x}(t_0,\mathbf{p}) = \mathbf{f}_0(\mathbf{p}), \tag{5}$$

$$\mathbf{p}\in P\subset\mathbb{R}^{n_p}. \tag{6}$$

This includes numerical optimal control problems for which the control input can be parameterized by a finite-dimensional parameter vector. A particular example is control and optimization of bioreactor models using dynamic flux balance analysis (DFBA) (Höffner, Harwood, & Barton, 2013; Mahadevan, Edwards, & Doyle, 2002). Derivative-based numerical methods to solve this optimization problem require, as the name suggests, derivatives of the cost function J with respect to the parameter vector $\mathbf{p}$. For differentiable problems, these can be evaluated via a solution of the forward sensitivity system (Feehery, Tolsma, & Barton, 1997; Hartman, 2002) or the adjoint system (Cao, Li, Petzold, & Serban, 2003). In this work the function h is locally Lipschitz continuous but not continuously differentiable. Therefore, the vector field $\mathbf{f}$ in Eq. (1) is also not necessarily continuously differentiable. For nondifferentiable optimization problems many complications arise such as the formulation of optimality conditions and the definition and computation of descent directions. One successful concept in this context is Clarke's generalized gradient (Clarke, 1990). In this work,

☆ The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Constantino M. Lagoa under the direction of Editor Richard Middleton.

E-mail addresses: hoeffner@mit.edu (K. Höffner), kamil@mit.edu (K.A. Khan), pib@mit.edu (P.I. Barton).

http://dx.doi.org/10.1016/j.automatica.2015.10.026
0005-1098/© 2015 Elsevier Ltd. All rights reserved.

an element of Clarke's generalized gradient is computed using Nesterov's lexicographic derivative (Nesterov, 2005). In addition to the nondifferentiability of the problem, the solution of (1) might not exist for all time $t \in [t_0, t_f]$ and therefore it is in general difficult to know *a priori* if the function J can be evaluated at $\mathbf{p}$.

1.1. Nonsmooth sensitivity analysis

The main contribution of this paper is the computation of an element of the generalized gradient of the objective function, which is sufficient information for many nonsmooth optimization methods. It presents the first result for the type of dynamic systems considered here. In Pang and Stewart (2009) the authors computed an element of a linear Newton approximation, which is a superset of the generalized gradient, for differential variational inequalities. A similar formulation for dynamic systems with a linear program embedded is conceivable. But, as discussed in Khan and Barton (2014), the linear Newton approximation does not satisfy certain properties, for example a linear Newton approximation containing the origin is not sufficient for global optimality of a convex function, and is therefore not suitable for our purposes. See Khan and Barton (2014) for illustrative examples.

Computation of an element of the generalized gradient is a difficult problem in practice. For piecewise differentiable functions with a factorable representation, methods based on automatic differentiation have been developed recently (Griewank, 2013; Khan & Barton, 2013). This approach is not practical for our purposes since explicit representation of the optimal solution value via parametric linear programming (Bemporad, Borrelli, & Morari, 2002) becomes intractable as the size of the linear program increases. In contrast, the approach presented here only requires the solution of a small number of linear programs.

It is known from parametric programming that the optimal solution value of a linear program as a function of the right hand-side of the constraints is a convex function and that the optimal dual solution set corresponds to the subdifferential. The following example illustrates that a linear Newton approximation based on an arbitrary choice of subgradient may not yield an element of the generalized gradient of the cost function.

Example 1.1. Consider the ODE[1]

$$\dot{x}(t,p) = (1-t)h(t,x(t,p)), \quad \forall t \in (0,2]$$
$$x(0,p) = p$$

with $p \in \mathbb{R}$ and

$$h(t,z) = \min_{\mathbf{v}\in\mathbb{R}^2} \; v_1 + v_2 \quad \text{s.t. } v_1 - v_2 = (1-t)z, \; \mathbf{v} \geq \mathbf{0}.$$

The unique solution of this ODE is given by

$$x(t,p) = \begin{cases} pe^{\frac{\operatorname{sign}(p)}{3}(1-(1-t)^3)} & \text{for } 0 \leq t < 1, \\ pe^{\frac{\operatorname{sign}(p)}{3}(1+(1-t)^3)} & \text{for } 1 \leq t \leq 2. \end{cases}$$

The map $\xi \equiv x(2,\cdot)$ is continuously differentiable for all $p \in \mathbb{R}$, hence the generalized gradient at $p = 0$ is given by $\partial\xi(0) = \{1\}$. Since an analytical solution rarely exists, in general we have to resort to numerical methods to evaluate an element of the generalized gradient. Computation of a linear Newton approximation is one approach, which is applied here. The dual problem of the embedded LP is

$$h(t,z) = \max_{\lambda\in\mathbb{R}} \; (1-t)z\lambda \quad \text{s.t.} \quad -1 \leq \lambda \leq 1.$$

Based on the results shown in the following section, it can be shown that a generalized gradient of the right-hand side function $f_t \equiv f(t,\cdot) = (1-t)h(t,\cdot)$ along the solution $x(\cdot,0)$ is given by

$$\partial f_t(x(t,0)) = \left\{(1-t)^2\lambda : \lambda \in [-1,1]\right\}, \quad \forall t \in (0,2]$$

and is non-singleton for almost all times. In contrast, the generalized gradient of ξ at $p = 0$ is a singleton. An element of a linear Newton approximation (Pang & Stewart, 2009) of $x(t,\cdot)$ at $p = 0$ is given by the solution of the ODE

$$\frac{d\Gamma x}{dt}(t,0) = A(t)\Gamma x(t,0), \quad \forall t \in (0,2],$$
$$\Gamma x(0,0) = 1,$$

for any Lebesgue integrable function $A : (0,2] \to \mathbb{R}$ such that $A(t) \in \partial f_t(x(t,0))$ for almost all $t \in (0,2]$. For $A(t) = \lambda(t)(1-t)^2$ with $\lambda(t) = 1$ for all t, it follows that $\Gamma x(2,0) = e^{1/3}$ which is not an element of the generalized gradient of ξ at $p = 0$. In contrast, if we choose $\lambda(t) = 1$ for $t < 1$ and $\lambda(t) = -1$ for $t \geq 1$, the solution of the above ODE system yields the unique element of the generalized gradient of ξ at $p = 0$. Hence, the choice of the element of the generalized gradient of the right-hand side is important. One interpretation of the result presented in this work is providing an approach which determines a correct element of the dual solution set for each time along a trajectory.

1.2. Dynamic flux balance analysis

Dynamic flux balance analysis provides a platform for detailed design, control and optimization of biochemical process technologies. It is a promising modeling framework that combines genome-scale metabolic network analysis with dynamic simulation of the extracellular environment. A DFBA model assumes that the intracellular fluxes are at equilibrium with the extracellular environment. The resulting underdetermined linear stoichiometric model is solved under the assumption of a cellular objective such as growth rate maximization. The model of the metabolism is coupled with the dynamic mass balance equations of the extracellular environment via expressions for the rates of substrate uptake and product excretion, which imposes additional constraints on the linear program defined by growth rate maximization of the cell. For example, simulation of batch or fed-batch yeast fermentation requires a description of the dynamics of the reactor, which are ODEs, and a model of the microbial agent represented by a linear program. The concentrations of the extracellular species evolve according to the dynamics of the differential equations. Meanwhile, the model of the microbial metabolism is used to predict the growth rate and production of ethanol and other metabolites that are of value, based on the consumption of sugars such as glucose and xylose from the extracellular environment. Existing optimization case studies using DFBA models in the literature do not address the nondifferentiable nature of the problem. Optimal control of genome-scale (more than 1000 LP variables) DFBA models has been limited due to the increased computation time for large models (Hjersted, Henson, & Mahadevan, 2007). In Hjersted and Henson (2006), Hjersted and Henson studied fed-batch optimization of a bioreactor with a small-scale model of yeast metabolism. The optimization problem was discretized temporally and the resulting mathematical program with equilibrium constraints (MPEC) was solved. This approach becomes intractable as the size of the FBA

[1] The ODE can be written in the form of (1)–(2) by adding a second state representing time, *i.e.*, $\dot{x}_2(t,p) = 1, x_2(0,p) = 0$, which is omitted to simplify the presentation.

Download English Version:

https://daneshyari.com/en/article/7109574

Download Persian Version:

https://daneshyari.com/article/7109574

Daneshyari.com