OPTIMIZATION AND CONTROL OF DISTRIBUTED PROCESSES

Matthias Heinkenschloss
Rice University, Houston, Texas, USA

Keywords: adjoint equation, adjoints, control, discretization, distributed processes, distributed systems, feasible set, Gauss-Newton method, gradient, Hessian, interior-point methods, Lagrange function, Lagrange multipliers, line-search method, Newton method, optimal control, optimal shape design, optimality conditions, parameter identification, quasi-Newton method, sensitivity equation, sequential quadratic programming, state, state equation, trust-region method

Contents

Summary

Distributed processes or systems involve spatially and, possibly, temporally-varying parameters. They are often modeled by partial differential equations (PDEs) or integral equations. The task of optimizing such a process arises when one tries to determine inputs into the process to make it more efficient or if one tries to identify system properties from measurements of inputs and their corresponding outputs. Optimization and control of distributed processes become increasingly more important because many

applications are modeled as distributed systems, and because recent developments in the numerical solution of PDEs, in optimal control theory, and in numerical optimization, as well as rapidly increasing computer power, make it now possible to solve such optimization problems at resolutions and within time frames appropriate for applications. The efficient and reliable solution of optimization problems governed by distributed processes requires the integration of theory and numerical solution of the underlying PDEs, of optimization and optimal control theory, and of numerical optimization.

This paper discusses the formulation of optimization problems governed by distributed processes, the characterization of their solutions, their discretization, and numerical methods for their solution. Optimization of distributed processes is an active research area. Most of the developments discussed in this paper took place in the last 40 years. Novel applications with an increased complexity generate new challenges that have motivated recent advancements and will fuel future developments in this area. Ever more powerful computer hardware and software enables the application of optimization methods to an increasing number of distributed processes and at the same time motivates new optimization algorithm development.

1. Introduction

In many science and engineering applications one seeks to influence the output of a given system through the application of suitable inputs, or one tries to identify system properties from measurements of inputs and their corresponding outputs. Some systems can be adequately described by finitely many, often time-dependent parameters. Such systems are called discrete or lumped parameter systems. Lumped parameter systems are modeled by (systems of) ordinary differential equations (ODEs) (see *Ordinary Differential Equations* and *Optimization of Ordinary Differential Equations*). Frequently, however, systems involve quantities that are distributed in space, and possibly in time. The spatial distribution of these quantities cannot adequately be represented by a fixed number of finitely many parameters. Such systems are called distributed parameter systems. They are modeled by partial differential equations (PDEs), integral equations, integrodifferential equations, or by more general functional equations (see *Partial Differential Equations*, *Integral Equations*, and *Integrodifferential Equations*). Inputs into the distributed system are represented by coefficients or coefficient functions in the PDE or by the shape of the domain in which the PDE is posed. The system output is obtained from the solution of the PDE. Examples of distributed parameter systems include the flow of air around an airfoil, the deformation of a plate due to an applied force, the flow of electric current through a body, the flow of water, oil, or gas through porous media, the temperature distribution inside a furnace, and the dependence of the value of a financial option on the price of the underlying asset. The distributed parameter systems in these examples are modeled by PDEs, which is assumed to be the case throughout. This simplifies our presentation. Most statements remain valid if a distributed parameter system is modeled by an integral equation, an integrodifferential equation, or by a more general functional equation.

The study of PDEs provides important insight into the phenomena they model. If a system can be modeled by a PDE, one can use computer simulations to predict its behavior. For example, the deformation of a plate due to an applied force can be modeled by a PDE. In the PDE model, the applied force determines the right-hand side of the PDE or its boundary conditions. The solution of the PDE—called the state—models the deformation of the plate. Material properties of the plate determine PDE coefficients. Thus by solving the PDE with varying parameters, one can estimate the deformation of a plate when exposed to varying forces, or one can predict the deformations of plates manufactured out of different materials when exposed to a certain force. Analogous procedures are used to simulate complex phenomena such as the flow of air around a cruising airplane, the impact of collisions on car bodies, the spread of contaminants in soil, and the flow of blood in a heart. Generally, PDE models are used to discover properties of the underlying system or they are used to modify the system so that it performs better. Single simulations are not sufficient to accomplish these tasks. Instead they lead to optimization problems in which objective function and constraint function evaluations involve the solution of the PDEs that model the system. Optimization problems may arise in the context of parameter estimation problems, optimal control problems or optimal shape design problems.

Optimization and control of distributed processes have increasingly become more important. This is because a growing number of applications require the resolution of system properties at a level that cannot be attained by simple algebraic models or lumped parameter models, but that requires PDE models. It is also due to recent developments in the numerical solution of PDEs, in optimal control theory, and in numerical optimization, as well as to rapidly increasing computer power, all of which now make it possible to solve such optimization problems at resolutions and within time frames appropriate for applications.

Optimization problems governed by distributed processes are challenging. The efficient and reliable solution of such problems requires the integration of the theory and the numerical solution of the underlying PDEs, of optimization and optimal control theory, and of numerical optimization (see *Partial Differential Equations, Numerical Solution of Partial Differential Equations, Optimization and Control of ODEs,* and *Nonlinear Programming.*)

2. Optimization Problems Governed by Distributed Processes

Abstractly, most optimization problems governed by distributed processes can be written in the language of functional analysis as

$$\begin{aligned} \min \quad & J(y,u), \\ s.t. \quad & c(y,u)=0, \\ & e(y,u)=0, \\ & h(y,u)\in K. \end{aligned} \tag{1}$$

In this abstract problem formulation, *u* denotes the control, the design parameter, or the parameter to be identified. The output of the system—called the state—is represented by *y*. These variables belong to infinite dimensional vector spaces *U* and *Y*, called the control space and the state space, respectively. The objective function $J: Y \times U \to R$, where R denotes the set of all real numbers, is a quantification of the performance measure. The variables y and u must be in the feasible set

$$F = \{ y \in Y, u \in U : c(y,u) = 0, e(y,u) = 0, h(y,u) \in K \}. \tag{2}$$

The constraints in (1) are described by three functions $c : Y \times U \to C$, $e : Y \times U \to E$, and $h : Y \times U \to H$, respectively, where C, E, and H are vector spaces. One distinguishes between two equality constraints. The first equality constraint $c(y,u) = 0$ represents the partial differential, integral, or functional equation that models the distributed process. This equation is called the governing equation or the state equation. The second equality constraint $e(y,u) = 0$, as well as $h(y,u) \in K$, where $K \subset H$ is a cone, represent auxiliary constraints on states and controls. The last constraint $h(y,u) \in K$ generalizes inequality constraints $h(y,u) \le 0$ in finite dimensions. The distinction between $c(y,u) = 0$ and $e(y,u) = 0$ is useful because these constraints are treated differently in several optimization algorithms for the solution of (1). To simplify the presentation, it is assumed that $c(y,u) = 0$ represents a PDE. In this paper u will be frequently called the control, even though it might actually represent a parameter to be identified or shape to be designed. The control space U, the state space Y, and the image spaces C, E, and H are assumed to be Banach spaces.

Most optimization problems governed by distributed processes fit into one of the following three classes of optimization problems:

- Parameter identification problems,
- Optimal control problems, and
- Shape optimization problems.

All of these problems can be formulated as optimization problems of the form (1). However, their problem characteristics are slightly different, and these differences can influence the choice of the optimization approach taken for their solution.

2.1. Parameter Identification Problems

Parameter identification problems, which are also known as parameter estimation problems or as inverse problems, arise when one tries to identify system properties from measurements of inputs and their corresponding outputs (see *Parameter Identification, Parameter Estimation, and Inverse Problems*). One application area that leads to parameter estimation problems is noninvasive material interrogation. Concrete examples include computer tomography and the identification of subsurface earth conductivity properties from electromagnetic measurements on the earth surface.

In the context of parameter identification problems, u denotes the parameter to be identified. The solution y of $c(y,u) = 0$ simulates the output of the system with parameter u. Given measurements y_i^{meas}, 1 = 1, …m**,** of the system output, one ideally wants to find u such that $c(y_i^{meas}, u) = 0$ for I = 1, …m. However, this problem is usually ill-posed. This means that a solution u does not exist, that the solution is not unique, or that small variations in the measurements y_i^{meas} lead to large variations in the solution u. A large amount of research in parameter identification is devoted to the development and analysis of so-called regularization techniques, that is, solution techniques that can cope with the ill-posedness and generate parameter estimates that approach the true parameter as measurement errors become small. Most regularization techniques are based on a formulation of the parameter identification problem as an optimization problem of the type (1). One typically tries to compute u so that the error between the given, measured output y_i^{meas} and the predicted output y (i.e,. the solution of $c(y,u) = 0$), is small. The objective function $J(y,u)$ is often a combination of a term that measures the error between y_i^{meas} and y and a so-called regularization term that introduces additional information about the parameter into the problem. Regularization terms may also be introduced through constraints of the form $h(y,u) \in K$. Details on regularization techniques for parameter identification problems are presented elsewhere (see *Parameter Identification, Parameter Estimation, and Inverse Problems*). Optimization problems (1) originating from parameter identification problems usually depend on some regularization parameter, which has to be carefully adjusted based on the computed solution y and u. One often has to solve sequences of optimization problems (1) to determine an appropriate regularization parameter.

2.2. Optimal Control Problems

In optimal control one wants to determine a system input u so that the performance of the system is optimized. If J is the performance measure, then this leads to a problem of the type (1). The constraint $c(y,u) = 0$ implicitly describes the dependence of system output y on the system input.

Unlike in parameter estimation there is no "true" system input that needs to be recovered, but one is able to choose among all technologically feasible inputs. Technological constraints on the system input u lead to constraints of the type $e(y,u) = 0$ or $h(y,u) \in K$. Such constraints can also arise when one tries to account for system disturbances that are unknown at the time the optimal control is computed.

In many applications, an approximate solution of (1) must be computed in a specified, short time frame. In such cases one cannot hope to compute a solution of (1), but only a control that achieves a sufficient improvement in system performance as measured by J.

-
-
-

Bibliography

Banks H.T. and Kunisch K. (1989). *Estimation Techniques for Distributed Parameter Systems*, 315 pp. Boston: Birkhäuser Verlag. [This monograph provides an introduction to approximation methods for least squares inverse problems governed by partial differential equations.]

Curtain R.F. and Zwart H.J. (1995). *An Introduction to Infinite Dimensional Linear Systems Theory*, 698 pp. Heidelberg: Springer Verlag. [This book provides an introduction to infinite dimensional systems with bounded inputs and outputs for which the state space is a Hilbert space.]

Fattorini H.O. (1999). *Infinite Dimensional Optimization and Control Theory*, 798 pp. Cambridge: Cambridge University Press. [This book studies existence results and necessary optimality conditions for optimal control problems governed by ordinary and partial differential equations.]

Fursikov A.V. (2000). *Optimal Control of Distributed Systems. Theory and Applications*, 305 pp. Providence, RI: American Mathematical Society. [This book studies existence results and necessary optimality conditions for optimal control problems governed by linear and nonlinear partial differential equations.]

Haslinger J. and Neittaanmäki P. (1996). *Finite Element Approximation for Optimal Shape, Material and Topology Design. Second edition*, 423 pp. Chicester: John Wiley & Sons. [This book presents theory, finite element discretization, and numerical solution of shape optimization, material, and topology design problems.]

Lions J.L. (1971). *Optimal Control of Systems Governed by Partial Differential Equations*, 396 pp. Heidelberg: Springer Verlag. [This is one of the first books on optimal control of partial differential equations and it provides a good overview of optimal control of linear partial differential equations.]

Nocedal J. and Wright S.J. (1999). *Numerical Optimization*, 636 pp. Heidelberg: Springer Verlag. [This is a recent book on methods for the solution of continuous finite dimensional nonlinear optimization problems.]

Sokolovski J. and Zolesio J.-P. (1992). *Introduction to Shape Optimization. Shape Sensitivity Analysis*, 250 pp. Heidelberg: Springer Verlag. [This book studies problem formulation, existence results, and shape sensitivity computations for shape optimization problems governed by linear partial differential equations.]

Biographical Sketch

Matthias Heinkenschloss is Associate Professor of Computational and Applied Mathematics in the George R. Brown School of Engineering at Rice University. He received his Doctoral degree in Mathematics from the University of Trier, Germany. He has done extensive research in the development, analysis, and implementation of algorithms for the solution of optimization problems governed by partial differential equations and their application to problems in science and engineering. Dr. Heinkenschloss's research papers have appeared in *ACM Transactions on Mathematical Software*, *Applied Mathematics and Computation*, *Control & Cybernetics*, *International Journal for Computational Fluid Mechanics*, *Journal of Optimization and Applications*, *Mathematical Programming*, *Optimization Methods and Software*, *SIAM Journal on Control and Optimization*, *SIAM Journal on Numerical Analysis*, and *SIAM Journal on Optimization.* He is on the editorial board of *Mathematical Programming*, *SIAM Journal on Control and Optimization*, *SIAM Review*, and *Systems and Control Letters*. Dr. Heinkenschloss is a member of the *Mathematical Programming Society* and the *Society for Industrial and Applied Mathematics (SIAM).*