Contributions in Mathematical and Computational Sciences 9

Thomas Carraro
Michael Geiger
Stefan Körkel
Rolf Rannacher *Editors*

Multiple Shooting and Time Domain Decomposition Methods

MuS-TDD, Heidelberg, May 6-8, 2013

MATCH HEIDELBERG UNIVERSITY

IWR

Springer

Contributions in Mathematical and Computational Sciences

Volume 9

More information about this series at http://www.springer.com/series/8861

Thomas Carraro • Michael Geiger •
Stefan Körkel • Rolf Rannacher
Editors

Multiple Shooting and Time Domain Decomposition Methods

MuS-TDD, Heidelberg, May 6-8, 2013

Springer

Editors
Thomas Carraro
Heidelberg University
Institute for Applied Mathematics
Heidelberg, Germany

Michael Geiger
Heidelberg University
Institute for Applied Mathematics
Heidelberg, Germany

Stefan Körkel
Heidelberg University
Interdisciplinary Center for Scientific Computing
Heidelberg, Germany

Rolf Rannacher
Heidelberg University
Institute for Applied Mathematics
Heidelberg, Germany

ISSN 2191-303X ISSN 2191-3048 (electronic)
Contributions in Mathematical and Computational Sciences
ISBN 978-3-319-23320-8 ISBN 978-3-319-23321-5 (eBook)
DOI 10.1007/978-3-319-23321-5

Library of Congress Control Number: 2015954722

Mathematics Subject Classification (2010): 34-XX, 34Bxx, 34B08k, 34B15, 34B60, 34Hxx, 34H05, 35-XX, 35D30, 35Jxx, 35Kxx, 35Q30, 35Q40, 35Q92, 49-XX, 49Mxx, 49M05, 49M15, 49M37, 62-XX, 62Kxx, 62K05, 62K25, 65-XX, 65Kxx, 65K05, 65K10, 65Lxx, 65L04, 65L05, 65L06, 5L09, 65L10, 65L80, 65Mxx, 65M20, 65M32, 65M60, 68-XX, 68Uxx, 68U20, 90-XX, 90Cxx, 90C30, 90C53, 90C55

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Preface to the Series

Contributions to Mathematical and Computational Sciences

Mathematical theories and methods and effective computational algorithms are crucial in coping with the challenges arising in the sciences and in many areas of their application. New concepts and approaches are necessary in order to overcome the complexity barriers particularly created by nonlinearity, high-dimensionality, multiple scales and uncertainty. Combining advanced mathematical and computational methods and computer technology is an essential key to achieving progress, often even in purely theoretical research.

The term mathematical sciences refers to mathematics and its genuine sub-fields, as well as to scientific disciplines that are based on mathematical concepts and methods, including sub-fields of the natural and life sciences, the engineering and social sciences and recently also of the humanities. It is a major aim of this series to integrate the different sub-fields within mathematics and the computational sciences and to build bridges to all academic disciplines, to industry and to other fields of society, where mathematical and computational methods are necessary tools for progress. Fundamental and application-oriented research will be covered in proper balance.

The series will further offer contributions on areas at the frontier of research, providing both detailed information on topical research and surveys of the state of the art in a manner not usually possible in standard journal publications. Its volumes are intended to cover themes involving more than just a single "spectral line" of the rich spectrum of mathematical and computational research.

The Mathematics Center Heidelberg (MATCH) and the Interdisciplinary Center for Scientific Computing (IWR) with its Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences (HGS) are in charge of providing and preparing the material for publication. A substantial part of the material will be acquired in workshops and symposia organized by these institutions in topical areas of research. The resulting volumes should be more than just proceedings collecting papers submitted in advance. The exchange of information and the discussions during the meetings should also have a substantial influence on the contributions.

Starting this series is a venture posing challenges to all partners involved. A unique style attracting a larger audience beyond the group experts in the subject areas of specific volumes will have to be developed.

The first volume covers the mathematics of knots in theory and application, a field that appears excellently suited for the start of the series. Furthermore, due to the role that famous mathematicians in Heidelberg like Herbert Seifert (1907–1996) played in the development of topology in general and knot theory in particular, Heidelberg seemed a fitting place to host the special activities underlying this volume.

Springer Verlag deserves our special appreciation for its most efficient support in structuring and initiating this series.

Heidelberg, Germany

Hans Georg Bock
Willi Jäger
Hans Knüpfer
Otmar Venjakob

Preface

The point of origin of this book was an international workshop with the same title (*Multiple Shooting and Time Domain Decomposition Methods—MuSTDD 2013*) that took place at the Interdisciplinary Center of Scientific Computing (IWR) at Heidelberg University in late spring 2013. Most of the chapters presented here are based on topics exposed in the talks given during this workshop.

The leading motivation for realizing this book project was its potential to fill a gap in the existing literature on time domain decomposition methods. So far, in contrast to domain decomposition methods for the spatial variables, which have found broad interest in the past two decades, the decomposition of the time domain still constitutes a niche. There is no comparable compendium on this subject, although an increasing amount of journal articles proves a growing need for these methods. Therefore, we firmly believe that this volume provides a useful overview over the state-of-the-art knowledge on the subject and offers a strong incentive for further research.

The book at hand is divided into two parts, which roughly reflect a classification of the articles into theoretical and application-oriented contributions:

- The first part comprises methodical, algorithmic, and implementational aspects of time domain decomposition methods. Although the context is often given by optimization problems (optimal control and parameter estimation with nonstationary differential equations), the covered topics are also accessible and crucial for researchers who intend to utilize time decomposition in a modeling and simulation framework. The topics covered in this theoretical part range from a historical survey of time domain decomposition methods via state-of-the-art environments for multiple shooting (such as ODE parameter estimation or DAE problems) up to recent research results, e.g. on different multiple shooting approaches for PDE, on multiple shooting in the optimal experimental design (OED) or the nonlinear model predictive control (NMPC) frameworks or on parareal methods as preconditioners.
- The second part is concerned with applications in different scientific areas that can potentially benefit from multiple shooting schemes and the related

parareal methods. In the application fields covered in this volume (amongst them fluid dynamics, data compression, image processing, computational biology, and fluid structure interaction problems), the two essential features of time domain decomposition methods, namely the stabilization of the solution process and its parallelizability, display their full potential.

Overall, we are convinced that this volume constitutes a unique compilation of methodical and application-oriented aspects of time domain decomposition useful for mathematicians, computer scientists, and researchers working in different application areas. Although it does not claim to be exhaustive, it provides a comprehensive accumulation of material that can both serve as a starting point for researchers who are interested in the subject and extend the horizon of experienced scientists who intend to deepen their knowledge.

We would like to acknowledge the support of several sponsors who made the MuSTDD workshop possible: the Priority Program 1253 of the German Research Association (DFG), the Mathematics Center Heidelberg (MATCH), and the Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences (HGS MathComp). Furthermore, we thank all the authors for their precious contributions. The cooperation with Springer, MATCH, and IWR should not be left unmentioned: it was a pleasure to work with them, and we thank all the people who rendered this 9th volume of *Contributions in Mathematical and Computational Sciences* possible by quietly and efficiently acting behind the scenes.

Heidelberg, Germany

Thomas Carraro
Michael Geiger
Stefan Körkel
Rolf Rannacher

Contents

Direct Multiple Shooting and Generalized Gauss-Newton Method for Parameter Estimation Problems in ODE Models

Hans Georg Bock, Ekaterina Kostina, and Johannes P. Schlöder

Abstract The paper presents a boundary value problem approach for optimization problems in nonlinear ordinary differential equations, in particular for parameter estimation, based on multiple shooting as originally introduced by Bock in the 1970s. A special emphasis is placed on the theoretical analysis including numerical stability, grid condition and a posteriori error analysis. The paper discusses advantages of multiple shooting versus single shooting which are illustrated by numerical examples.

1 Introduction

The history of shooting methods for optimization problems in differential equations goes back to the 1950s when shooting methods were first used to solve two point boundary value problems (TPBVP) resulting from application of the Pontryagin maximum principle to optimal control problems, see e.g. [26, 31, 38, 40]. A first versatile algorithm capable to treat TPBVP with switching points has been developed by Bulirsch and Stoer [18] giving start to numerous theses in the Bulirsch group extending optimal control theory and multiple shooting algorithms for TPBVP. The drawbacks of this "indirect" approach are that the boundary value problems resulting from the maximum principle for optimal control problems are difficult to derive, moreover, they are usually ill-conditioned and highly nonlinear in terms of state and adjoint variables, they have jumps and switching conditions.

Another type of methods for optimal control problems—"direct" single shooting methods—appeared in the 1960s. Instead of using the maximum principle and adjoint variables, in these methods the controls were discretized, such that solving differential equations resulted in finite nonlinear programming problems. Numerous

H.G. Bock • J.P. Schlöder
IWR, Heidelberg University, Heidelberg, Germany
e-mail: bock@iwr.uni-heidelberg.de; schloeder@iwr.uni-heidelberg.de

E. Kostina (✉)
IAM, Heidelberg University, Heidelberg, Germany
e-mail: kostina@uni-heidelberg.de

T. Carraro et al. (eds.), *Multiple Shooting and Time Domain Decomposition Methods*, Contributions in Mathematical and Computational Sciences 9,
DOI 10.1007/978-3-319-23321-5_1

direct algorithms were developed, usually as feasible step gradient type methods, e.g. [17, 36, 37]. The reason for this was that effective non-feasible step algorithms for constrained nonlinear optimization emerged only in the 1970s, i.e. SQP [25, 41].

In this paper we consider a "direct" multiple shooting method for optimization problems with differential equations, also known nowadays as "optimization boundary value problem", "all-at-once", or "simultaneous" approaches. This approach is based on constrained nonlinear optimization instead of the maximum principle, and the discretized BVP is treated as an equality constraint in the optimization problem which is then solved by non-feasible step methods. This approach became a standard tool for solving optimization problems for differential equation models, see [9, 10, 12, 13].

The multiple shooting combined with the generalized Gauss-Newton is especially suitable for parameter estimation.

In this paper we focus on the multiple shooting for parameter estimation for differential equations. A special emphasis is placed on the theoretical analysis including numerical stability, grid condition and a posteriori error analysis. We discuss also advantages of multiple shooting versus single shooting which are illustrated by numerical examples.

2 Parameter Estimation Problem in ODE Systems

We consider parameter estimation problems that are characterized by a system of ordinary differential equations for state variables $x(t)$

$$\dot{x} = f(t, x, p),$$

the right hand side of which depends on a parameter vector p. Furthermore, measurements η_{ij} for the state variables or more general for functions in the states are given

$$\eta_{ij} = g_{ij}(x(t_j), p) + \varepsilon_{ij},$$

which are collected at measurement times $t_j, j = 1, \ldots, k$,

$$t_0 \le t_1 < \ldots < t_k \le t_f,$$

over a period $[t_0, t_f]$, and are assumed to be affected by a measurement error ε_{ij}.

The unknown parameters p have to be determined such that the measured (observed) process is "optimally" reproduced by the model. If the measurement errors ε_{ij} are independent, Gaussian with zero mean value and variances σ_{ij}^2 an appropriate objective function is given by a weighted l_2-norm of the measurement errors

$$l_2(x, p) := \sum_{i,j} \sigma_{ij}^{-2} \varepsilon_{ij}^2 = \sum_{i,j} \sigma_{ij}^{-2} (\eta_{ij} - g_{ij}(x(t_j), p))^2. \quad (1)$$

In this case the minimization of the weighted squared errors provides a maximum-likelihood estimator for the unknown parameter vector.

In a more general case where the measurement errors are correlated with a known (positive definite) covariance matrix C, we have to use C^{-1} as a weight for the definition of a scalar product replacing (1) to receive a maximum-likelihood estimator.

In many problems additional (point-wise) equality and/or inequality constraints on parameters and state variables arise as restrictions onto the model

$$e(t_j, x(t_j), p) = 0,\ u(t_j, x(t_j), p) \geq 0, j = 1, \ldots, k.$$

For notation simplicity we assume the constraints are stated at the same points as measurements.

A rather general parameter estimation problem for ordinary differential equations can be summarized as follows:

Problem [PE1] Find a *parameter vector* $p \in \mathbb{R}^{n_p}$ and a trajectory $x : [t_0, t_f] \to \mathbb{R}^{n_d}$, that minimize the objective function (describing a weighted norm of measurement errors)

$$l_2 = \|r_1(x(t_1), \ldots, x(t_k), p)\|_2, r_1 \in \mathbb{R}^{n_1}, \tag{2}$$

where n_1 is a number of all measurements. The solution has to satisfy the *system of ordinary differential equations* of dimension n_d

$$\dot{x} = f(t, x, p), \quad t \in [t_0, t_f], \tag{3}$$

the n_2 *equality constraints*

$$r_2(x(t_1), \ldots, x(t_k), p) = 0 \tag{4}$$

and the n_3 *inequality constraints*

$$r_3(x(t_1), \ldots, x(t_k), p) \geq 0. \tag{5}$$

Other functionals than the l_2-norm can be used, like the general l_q-functional

$$l_q(x, p) = \sum_{i,j} \left|\beta_{ij}(\eta_{ij} - g_{ij}(x(t_j), p))\right|^q,\ 1 \leq q < \infty$$

with the sum of absolute values of errors for $q = 1$ as the most significant special case besides $q = 2$, and the limit case of the Chebyshev or minimax problem

$$l_\infty(x, p) = \max_{i,j} \left|\gamma_{ij}^{-1}(\eta_{ij} - g_{ij}(x(t_j), p))\right|.$$

l_1- and l_∞-estimators have some specific properties which make them interesting. Under certain regularity assumptions l_∞-optimization leads to a solution in which exactly $n+1$ of the weighted measurement errors take the maximum value, whereas exactly n errors vanish in case of the l_1-optimization so that the l_1-optimal solution interpolates n measurement values. Here $n = n_v - n_{\mathrm{con}}$ is the number of the remaining degrees of freedom, where $n_v = n_d + n_p$ counts the differential equations and parameters, n_{con} counts the equality and active inequality constraints at the solution. Hence, the optimal solution is only specified by few "best" (l_1) and "worst" (l_∞) measurement values, respectively. From statistical point of view l_∞-based estimator provides a maximum-likelihood estimation for uniformly distributed errors with maximum γ_{ij}, l_q-based estimator is related to distributions of the type $\alpha_{ij} \exp(-|\beta_{ij}\varepsilon_{ij}|^q)$. For further discussion see [2, 15, 32].

3 Initial Value Problem Approach

The simplest—and maybe most obvious—approach for the numerical treatment of parameter estimation problems in differential equations is the repeated *solution of the initial value problem* (IVP) for fixed parameter values in framework of an iterative procedure for *refinement of the parameters* to improve the parameter estimates and to fulfill possible constraints on states and parameters. Thus, the inverse problem is lead back to a sequence of IVPs.

Besides the undeniable advantage of a simple implementability the IVP approach has two severe fundamental disadvantages which are clearly shown in numerical practice and are verified by theoretical analysis.

On the one hand the state variables $x(t)$ are eliminated—by means of differential equation (3)—in favor of the unknown parameters p by the re-inversion of the inverse problem. As a consequence any information during the solution process that is especially characteristic for the inverse problem is disregarded. Consequently, the structure of the inverse problem is destroyed.

On the other hand the elimination of the state variables can cause a drastical loss of stability of the numerical procedure. At least for bad initial guesses of the parameters, which always have to be expected in practice, the (non-linear) initial value problem can be ill-conditioned and difficult to solve or can be not solvable at all even if the inverse problem is well-conditioned. As a consequence the IVP approach places high demands on the used iterative method or on the quality of the initial values.

Let us illustrate these properties of the IVP approach by two examples.