Mathematisches Forschungsinstitut Oberwolfach

Report No. 3/2004

Numerical Methods for Instationary Control Problems

Organised by
Karl Kunisch (Graz)
Angela Kunoth (Bonn)
Rolf Rannacher (Heidelberg)

January 18th – January 24th, 2004

Introduction by the Organisers

The topic for the current Mini–Workshop organized by Karl Kunisch (Univ. Graz), Angela Kunoth (Univ. Bonn) and Rolf Rannacher (Univ. Heidelberg) emerged from the Oberwolfach Workshop "Numerical Techniques for Optimization Problems with PDE Constraints" which was held February 16–22, 2003. It was realized that numerically solving control problems which are constrained by time–dependent nonlinear PDEs (Partial Differential Equations) are particularly challenging with respect to the complexity of the problem.

Mathematically one has to minimize a functional under PDE constraints and possibly additional constraints on the state and the control. Standard discretizations on uniform grids in space and time will only yield solutions where the inherent structures of the problem (nonlinearity, constraints) are not sufficiently captured. Certain optimization problems for large coupled systems of partial differential equations are currently not yet numerically treatable or do not satisfy the time constraints required in practice. Overcoming this barrier can only be achieved by designing new mathematically founded algorithmic approaches. The road towards this goal leads to many interesting problems in optimization, linear algebra, numerics, analysis, and approximation theory.

The conference had 21 participants which represented continuous optimization, numerical analysis and scientific computing. 18 talks were given. The 10 longer, overview–style talks were on optimization with PDEs, focussing on

- Modelling and Global Optimization
- Snapshot Selection
- Treatment of State Constraints

- ODE Techniques for PDEs
- Automatic Differentiation
- Adaptive Finite Elements
- Parameter Estimation
- Adaptive Wavelets.

These talks were intended to bridge the gap between the different research fields Optimization and Numerics. They were complemented by 8 shorter talks on more specialized research topics, ranging from efficiency indices for optimization over iterative methods for the coupled systems and multigrid acceleration to modelling issues in optimization, and data compression by means of proper orthogonal decompositions and central Voronoi tesselations.

Different modern approaches to overcome the complexity issues in numerical simulations for PDE–constrained optimization have been presented and discussed. One of the approaches is to employ fast iterative solvers like multigrid on uniform grids. The methodology which conceptually provides the largest potential is to introduce adaptivity. This drastically reduces complexity but depending on the context may require solving an additional problem. Wavelet approaches particularly allow to resolve each of the variables separately and in addition provide a built–in preconditioning. Yet another approach uses compressed information in order to efficiently solve the primal–dual system.

All these issues are currently very active research areas. The extensive discussions held during this workshop have produced a number of new ideas and connections. It was agreed upon that a mere black–box–style matching of efficient PDE codes with optimization tool boxes would on one hand remain much below its potential and on the other hand not help overcoming complexity barriers. Some concepts from automatic differentiation seem to carry over to adaptive methods. Even combining adaptivity with proper orthogonal decompositions may be a very promising direction. The many new ideas discussed during this workshop will have to be further elaborated in future.

Mini–Workshop on Numerical Methods for Instationary Control Problems

Table of Contents

Abstracts

Adaptivity for Optimization of Time-dependent Partial Differential Equations

Roland Becker

In this talk, we derive a posteriori error estimators for optimal control problems. The error estimates are general and apply to Galerkin discretizations of optimization problems governed by partial differential equations. In the first part of the talk we describe some algorithmic aspects of a resulting adaptive algorithm. We focus on the case of time-dependent partial differential equations, where we want to adapt the step size and the dynamically varying meshes for space discretization. This estimate can be used to derive automatic adaptation of $(h, p) \times (k, r)$ method where h denotes the spatial mesh, p the distribution of polynomial degree in the spatial mesh h, k the temporal mesh, and r the polynomial degree in time. In order to simplify the situation we focus on the case where only the mesh h is to be adapted dynamically and the control is frozen, since this seems to be the difficult part. The goal of our adaptive algorithm is to find a method which has computing time linear with respect to the overall number of unknowns and storage requirements proportional to the temporal mean of the employed meshes. In order to achieve this goal, we have to use a divide-and-conquer algorithm as the checkpointing/windowing algorithm known from automatic differentiation and optimal control. The essential additional difficulty in our context is the fact that we need information about the co-state (solving a backwards-in-time equation) when computing forward. This is due to the structure of the error estimator.

The second part of the talk describes a posteriori error estimates for optimization problems. We consider the following general case: we are interest in computing an interest function $I(q, u)$ which depends on both control q and state variable u. The interest functional is independent of the optimization problem which determines q and u. By specialization we obtain estimators for:

- error in the cost functional [2, 3]
- error in a functional of the controls [4]
- error in an independent functional of the state [5]
- norm of the error of controls [1]

Beside the first estimator, the others require the solution of an additional problem involving the adjoint of the linearized state operator. The right-hand side of this problem depends on the special context.

In the last part of the talk, we show that the information given by the additional adjoint problem might be used for further purposes. We employ the well-known concept of condition numbers in order to produce answers to the following questions: which parameters have been most important, which measurements have been most important in computing the solution?

References

[1] R. Becker, Adaptive finite element methods for optimal control of partial differential equations: Norms of the error, Technical report, LMA, Université de Pau et du Pays de l'Adour, 2004.
[2] R. Becker, H. Kapp, R. Rannacher, Adaptive finite element methods for optimal control of partial differential equations: Basic concepts, SIAM J. Contr. Optim., 39(1):113–132, 2000.
[3] R. Becker, R. Rannacher, An optimal control approach to a-posteriori error estimation, Acta Numerica 2001, pp. 1–101, Cambridge University Press.
[4] R. Becker, B. Vexler, A posteriori error estimation for finite element discretizations of parameter identification problems, to appear in Numer. Math.
[5] R. Becker, B. Vexler, Numerical sensitivity analysis with adaptive finite elements: Applications to CFD problems, submitted.

Fast Integral Equation Solvers and Applications to Problems with Dynamic Interfaces

George Biros

(joint work with L. Ying and D. Zorin (New York University))

Boundary integral formulations have been extensively used for the analysis and numerical solution of elliptic partial differential equations. However, with the exception of problems in inverse scattering, there has been limited work in boundary integral equation formulations for optimal control and optimal design problems. There are several reasons for that: prohibitive complexity of efficient implementations for non-Laplacian kernels, difficulty with distributed force terms, and restriction to problems with piecewise constant coefficients.

Recent developments however, indicate integral equation formulations might have impact to a larger class of problems. We present two new algorithms and two examples that illustrate the efficiency of the new methods: (1) A fast solver for Stokes and Navier-Stokes equations, (2) a new kernel-independent fast-multipole method for kernels related to constant coefficient elliptic PDEs, (3) a three-dimensional rigid body-fluid interaction problem, and (4) a prototypical shape optimization problem of a Dirichlet interior Stokes problem.

1. Fast Solvers for Stokes and Navier-Stokes Equations. Our motivation in designing this method is the design of efficient algorithms for flows with dynamic interfaces. Solvers for such problems should be built on algorithms that do not require expensive preprocessing, like unstructured mesh generation, since the interface is moved to a new location at each time step. The main features of the method are: It can be applied to arbitrary piecewise smooth geometries; It does not require mesh generation; It can solve problems with distributed forces; It is second-order accurate and readily generalizes to arbitrary order of accuracy. If an optimal boundary integral equation solver is used, the method has $\mathcal{O}(N)$ complexity.

Our method is based on potential theory for linear second-order elliptic operators. Using an indirect integral formulation, the solution of a Dirichlet problem

can be written as the sum of a double layer potential and a *Newton potential* (the domain convolution of the Green's function with the distributed force). Under such a scheme the evaluation of the solution must consist of three steps: (1) computation of the Newton potential, (2) solution of a boundary integral equation to compute a double layer potential that satisfies the boundary conditions, and (3) evaluation of a double layer potential everywhere in the domain. Details can be found in [2] for the Stokes equations and [1] for the unsteady Navier-Stokes problem.

2. Kernel-independent Fast Multipole Method. The main feature of the new method is its black-box application to several different non-oscillatory elliptic kernels. Our algorithm has the structure of the adaptive FMM algorithm [3] but requires only kernel evaluations without sacrificing accuracy and efficiency compared to the original algorithm. The crucial element of our approach is to replace the analytic expansions and translations with *equivalent density* representations. These representations are computed by solving local exterior and interior problems on circles (2D), spheres or cubes (3D) using integral equation formulations. We have demonstrated the efficiency of our method in both 2D and 3D for many kernels: the single and double layer potentials of the Laplacian, the modified Laplacian, the Navier, the Stokes, and their modified variants. Our method has $\mathcal{O}(N \log N)$ asymptotic complexity, whereas for reasonable assumptions on the initial particle distribution the complexity becomes $\mathcal{O}(N)$. Like analytic FMM, our method works well for nonuniform particle distributions. Details can be found in [4]. We have also developed an MPI-based parallel version of the method, and have performed systematic scalability tests. Overall we have achieved very good iso-granular and fixed-size scalability on up to 3000 processors. A detailed discussion can be found in [5].

3. Fluid-structure Interaction Formulation. We have developed algorithms to simulate the interaction of rigid bodies of arbitrary geometry with Stokesian fluids, ignoring inertial terms in the fluid and using an integral formulation for the equations which describe the motion of the dynamics. These equations are a set of integrodifferential equations the interaction between the fluid and a rigid object and consist of the linear and angular momentum conservation for the rigid body and the Stokes equations for the fluid. The coupling is induced by the requirement of non-slip condition and force balance across the interface. These results are work in progress. For the fluid-structure interaction runs we are currently working on convergence studies, and we have not performed systematic scalability analysis. However, the main cost in these simulations is the solution of the underlying integral equations. To this end we are working on efficient preconditioning schemes.

4. Shape Optimization for Stokes-constrained Systems. We present a 2D shape optimization problem for the interior Dirichlet problem of a Stokesian fluid. The flow is represented using a second kind integral equation formulation. The objective function is of tracking type. The boundary is represented using a periodic B-spline and the optimization variables are the control points. Adjoints

are used to compute shape sensitivities and the problem is solved using a reduced space quasi-Newton method globalized by trust-region. The first derivatives of the adjoint and forward problem involve hypersingular kernels (Hilbert transforms) which are approximated using an odd-even Nyström integration scheme. The results, although very preliminary, are very encouraging since only a small number of quasi-Newton iterations are sufficient for convergence.

References

[1] G. Biros, L. Ying, and D. Zorin, The embedded boundary integral method for the unsteady incompressible Navier-Stokes equations, Technical Report TR2003-838, Courant Institute, New York University, 2002.
[2] G. Biros, L. Ying, and D. Zorin, A fast solver for the Stokes equations with distributed forces in complex geometries, Journal of Computational Physics, 193(1):317–348, 2003.
[3] L. Greengard, The Rapid Evaluation of Potential Fields in Particle Systems, MIT Press, Cambridge, MA, 1988.
[4] G. Biros, L. Ying, and D. Zorin, A kernel-independent adaptive fast multipole method in two and three dimensions, Journal of Computational Physics, to appear.
[5] G. Biros, L. Ying, D. Zorin, and H. Langston, A new parallel kernel-independent fast multiple algorithm, in: Proceedings of SC03, The SCxy Conference series, Phoenix, Arizona, November 2003, ACM/IEEE.

On Space-time Multigrid Solution of Unsteady Optimal Control Problems

Alfio Borzì
(joint work with K. Kunisch and R. Griesse)

The development and investigation of space-time multigrid schemes for unsteady reaction diffusion optimal control problems are reported. We focus on the control of the time evolution of chemical and biological processes characterized by non-monotone nonlinearities. For benchmarking our algorithms, we propose two models:

The solid fuel ignition model

$$-\partial_t y + \delta e^y + \Delta y = u, \quad \delta > 0,$$

results in a singular optimal control problem which cannot be solved by any method based on a free evolution of the state variable.

The lambda-omega system is given by

$$\frac{\partial}{\partial t}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{bmatrix} \lambda(y_1,y_2) & -\omega(y_1,y_2) \\ \omega(y_1,y_2) & \lambda(y_1,y_2) \end{bmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \sigma\,\Delta\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

where

$$\lambda(y_1,y_2) = 1 - (y_1^2 + y_2^2) \quad \text{and} \quad \omega(y_1,y_2) = -\beta(y_1^2 + y_2^2).$$

The functional form of λ and ω was proposed in [4] to model chemical turbulence. The evolution into a chaotic state of the $\lambda - \omega$ system can also be observed from a principal component analysis via proper orthogonal decomposition of its

snapshots. As the system becomes less ordered, the energy content becomes more and more evenly distributed among the eigenmodes.

In the first model case, the control, represented by u, is applied to avoid blow-up or to optimize the combustion process [1]. In the second case the control (u_1, u_2) is applied to drive the system from a turbulent to a regular state [2]. The optimality systems characterizing the optimal control solution are solved by space-time multigrid schemes with typical multigrid efficiency and robustness with respect to the choice of the optimization parameters. These features are obtained by developing appropriate collective smoothing schemes.

Using two-grid Fourier analysis, sharp estimates of convergence factors are obtained for linear model problems [3]. Results of numerical experiments demonstrate that these estimates remain sharp also for the nonlinear cases considered here.

References

[1] A. Borzì, K. Kunisch, A multigrid method for optimal control of time-dependent reaction diffusion processes, In K.H. Hoffmann et al. (eds.), Fast Solution of Discretized Optimization Problems, Internat. Ser. on Numerical Mathematics, Vol. 138, Birkhäuser, 2001.

[2] A. Borzì, R. Griesse, Distributed optimal control of lambda-omega systems, submitted to SICON.

[3] A. Borzì, Multigrid methods for parabolic distributed optimal control problems, J. Comp. Appl. Math, 157 (2003), pp. 365–382.

[4] Y. Kuramoto, S. Koga, Turbulized rotating chemical waves, Prog. Theor. Phys., 66(3) (1981), pp. 1081–1085.

Algorithms for Optimization of Time-dependent PDE Systems: Can We Realize the Same Efficiencies as in the Steady Case?

Omar Ghattas

(joint work with G. Biros (University of Pennsylvania), V. Akcelik, J. Bielak, and I. Epanomeritakis(Carnegie Mellon University))

The answer to the question posed in the title depends of course on the type of problem being solved. We begin by recalling some of our earlier work on fast solvers for optimization problems that are governed by PDEs [1,2,3,4,5]. The method we developed, which we refer to as *Lagrange-Newton-Krylov-Schur (LNKS)*, solves the full optimality system consisting of state, adjoint, and control equations using an inexact preconditioned Newton-QMR method. The preconditioner is a block factorization that emulates a reduced quasi-Newton SQP method: it approximates the reduced Hessian via suitably-initialized limited memory BFGS updates while discarding other second derivative terms, and replaces the exact state and adjoint solves with application of appropriate preconditioners, e.g. additive Schwarz or multigrid. If sufficient descent cannot be obtained with a line search, then we drop down to the reduced space and take a quasi-Newton step. Experiments with this

method on some problems of optimal control of three-dimensional steady Navier-Stokes flows via boundary suction/injection demonstrate high parallel scalability, mesh-independence of Newton iterations, mesh-independence of Krylov iterations (provided an optimal state preconditioner is available), and solution to optimality in four times the cost of a flow solution, for a problem with over 600,000 state and 9000 control variables. This small constant multiple of the state solve cost is due to iterating in the full space, which folds the iterations (linear and nonlinear) needed to converge the flow into those required for optimization. LNKS is most effective when the state equations are difficult to solve, requiring many iterations.

We next discuss an inverse parameter estimation problem governed by earthquake modeling via the elastic wave equation. The problem is to find the distribution of elastic parameters of large sedimentary basins such as Los Angeles, from surface observations of past earthquakes [6,7]. The forward problem alone requires terascale computing: our typical earthquake simulations resolve up to 1 Hz ground motion frequencies, involve 100 million grid points and 40,000 times steps, and require several hours of run time on the 3000 processor AlphaCluster supercomputer in Pittsburgh. The inverse problem is formulated via output least squares, regularized by a total variation (TV) functional. TV eliminates oscillatory components of the material properties, while preserving discontinuities at material interfaces. The solver is an inexact Newton-CG method in the reduced space, with the same preconditioning as in LNKS. However, because we iterate in the reduced space, an exact forward and adjoint solve are required at each CG iteration. Mesh independence of Newton and CG iterations is observed, and the number of inner and outer iterations required for convergence is similar to those observed for the flow control problem. However, the difference here is the requirement for exact solution of the forward and adjoint wave equation. For the largest inverse problem we solved, involving 17 million inversion parameters, the product of inner and outer iterations is such that 800 total forward and adjoint wave propagations are required. The essentially renders the inverse problem intractable for our goal of reconstructing the structure of the LA Basin to a 1 Hz frequency resolution.

We conclude the talk with a somewhat pessimistic discussion of several opportunities for speeding up the convergence of the earthquake inversion problem. Additional processors won't help, since the granularity of the computation is low to begin with. We could switch from a reduced space solver to a full space LNKS method, but there is nothing to be gained since the explicit forward solver offers little opportunity for approximation with a simpler solve. A coarser mesh misses the finest wavelengths, and a longer time step loses the shorter periods, both of which contribute importantly to the surface response. Implicit methods are not useful for wave propagation problems in which the system is responding in all of its resolvable scales. Adding processors in the time direction is not helpful, since information propagates at finite speed. Reduced order modeling in the state space is likely unproductive, since the system is responding at all scales (indeed the mesh was designed to just resolve the finest scales of interest). Similarly, reduced order modeling in the parameter space faces the problem of trying to construct a response

surface in a very high dimensional (e.g. 10 million) space. Some improvements in the linear preconditioner can probably be made by exploiting the compact and differential structure of the reduced Hessian, but since we have typically just 20 CG iterations per Newton iteration, the reduction in iterations must be balanced against the cost of construction of the preconditioners. Similar statements can be made about improvements in the nonlinear solver, e.g. through nonlinear preconditioning. We conclude that the inverse earthquake modeling problem to frequencies of engineering interest remains a major challenge, when measured against the cost of the forward simulation.

References

[1] G. Biros, O. Ghattas, Parallel domain decomposition methods for optimal control of viscous incompressible flows, Proceedings of Parallel CFD '99, Williamsburg, Virginia, USA, May 1999.

[2] G. Biros, O. Ghattas, Parallel Newton-Krylov methods for PDE-constrained optimization, Proceedings of SC1999, IEEE/ACM, Portland, OR, USA, November 1999.

[3] G. Biros, Parallel Lagrange-Newton-Krylov-Schur Methods for PDE-constrained Optimization, with Application to Optimal Control of Viscous Flows, Ph.D. Thesis, Carnegie Mellon University, September 2000.

[4] G. Biros and O. Ghattas, Parallel Lagrange-Newton-Krylov-Schur methods for PDE-constrained optimization. Part I: The Krylov-Schur solver, to appear in SIAM J. on Scientific Computing.

[5] G. Biros and O. Ghattas, Parallel Lagrange-Newton-Krylov-Schur methods for PDE-constrained optimization. Part II: The Lagrange-Newton solver, and its application to optimal control of steady viscous flows, to appear in SIAM J. on Scientific Computing.

[6] V. Akcelik, G. Biros, and O. Ghattas, Parallel multiscale Gauss-Newton-Krylov methods for inverse wave propagation, Proceedings of SC2002, Baltimore, MD, USA, November 2002.

[7] V. Akcelik, J. Bielak, G. Biros, I. Epanomeritakis, A. Fernandez, O. Ghattas, E. Kim, J. Lopez, D. O'Hallaron, T. Tu, and J. Urbanic, High-resolution forward and inverse earthquake modeling on terascale computers, Proceedings of SC2003, IEEE/ACM, Phoenix, AZ, USA, November 2003.

From Algorithmic Differentiation to Automated Design

Andreas Griewank

To facilitate the transition from simulation to optimization we suppose that we are provided (only) with an iterative solver for some state equation and a procedure for evaluating an objective function. This is a realistic scenario in aerodynamics, where the state equation is some discretized variant of the Navier Stokes equation governing the flow around a wing, and the objective may be the drag, which is to be minimized by varying the design of the wing. Since the state space may have very high dimensionality we prefer not to modify the given solver but merely assume that it is contractive as an iterative map. However, using algorithmic differentiation we may derive from it, in an automated fashion, a fixed point solver for the corresponding adjoint equation and the computation of an approximate reduced gradient. This methodology had been introduced by Griewank and Faure under the name of piggy-back differentiation [1].

In this talk we show first that while the two fixed point iterations converge with the same asymptotic contractivity rate, the accuracy of succesive approximate solutions to the adjoint equation lags behind that for the underlying state space iterates. More specifically, the ratio between the residuals of the two equations at the k-th, coupled iteration grows linearly with k [2].

Secondly, we examine the choice of a matrix for preconditioning the approximate reduced gradient in a simultaneous update of the design variables. As it turns out the seemingly natural choice of the reduced Hessian is not the best choice, but may lead to divergence. Instead we find that local convergence can be assured by projecting the Lagrangian Hessian onto another subspace, at least when the full Hessian is positive definite [3]. These theoretical observation are confirmed numerically on a 2D test problem provided by Volker Schulz. Either of the two projected Hessians can also be evaluated by automatic differentiation, so that we obtain a methodology for optimal design in a fairly automated fashion. Practical validations on Euler and Navier Stokes codes for 3D and 2D flows are under way.

References

[1] A. Griewank, C. Faure, Reduced functions, gradients and hessians from fixed point iteration for state equations, Numer. Algor., 30 (2002), no. 2, 113–139.
[2] A. Griewank, C. Faure, Linear and quadratic delay of reduced gradients and Hessians obtained from fixed point iterations for state equations, in preparation.
[3] A. Griewank, The correct projected Hessian for preconditioning the approximate reduced gradient in piggy-back optimization, in preparation.

Design of Experiments and Snapshot Generation in Reduced-order Modeling

Max Gunzburger

Reduced-order modeling strategies such as proper orthogonal decomposition (POD) methods are developed from a set of snapshots. The reduced-order model cannot contain more information than that contained in the snapshot set. Thus, the generation of snapshots is crucial to the success of reduced-order models. The generation of snapshots is an exercise in the design of experiments, i.e., how does one choose the values of the parameters used to generate the snapshot simulations or the time instants at which one evaluates the snapshots? We discuss the use of design of experiment-based strategies for parameter selection for snapshot generation. Issues that arise in selecting a method for sampling points in parameter space are considered, and the relative merits of different methods (e.g., quasi-Monte Carlo sequences, Latin hypercube sampling, centroidal Voronoi tessellation, etc.) are discussed. Several notions of uniformity for point sets are compared, as are their effect on deciding which sampling methods are best for specific applications. Also, the role of known information about the parameters and how to incorporate this information into the point sampling process for snapshot generation are considered.

A major consideration of the talk is a new point sampling strategy that we have developed that is based on centroidal Voronoi tessellations (CVT's). These are special Voronoi tessellations for which the generators of the Voronoi tessellation are also the centers of mass, with respect to a given density function, of the associated Voronoi cells. CVT's have many uses in many applications; in particular, CVT's can be used for generating very high-quality point sets in regions and on surfaces. Using several volumetric measures of uniformity, CVT point samples are shown to be more "uniform" that those obtained by existing strategies. On the other hand, CVT point sets do not have good properties when projected onto lower dimensional surfaces, e.g., the faces of a hypercube. Such a property is desirable in some applications such as high-dimensional integration. For the latter application, one can define "Latinized" CVT point sets that possess good projections. For design of experiment applications relevant to snapshot generation, both CVT and Latinized CVT point sets are superior to existing points sampling methods. Incidentally, CVT strategies also offer an alternative to POD as a means for defining a reduced basis from a set of snapshots.

Various aspects of the talk represent joint work with John Burkardt, Hoa Nguyen, Janet Peterson, and Yuki Saka (Florida State University) and Hyung-Chun Lee (Ajou University, Korea).

References

[1] J. Burkardt, M. Gunzburger, and H.-C. Lee, Centroidal Voronoi tessellation-based reduced-order modeling of complex systems, to appear in SIAM J. on Scientific Computing.

[2] J. Burkardt, M. Gunzburger, H. Nguyen, J. Peterson, and Y. Saka,Centroidal Voronoi tessellation points sampling I: Uniform sampling in hypercubes, in preparation.

[3] J. Burkardt, M. Gunzburger, H. Nguyen, J. Peterson, and Y. Saka, Centroidal Voronoi tessellation point sampling II: Nonuniform, anisotropic sampling in general regions, in preparation.

Adjoint-based Adaptive Time-stepping for Partial Differential Equations using Proper Orthogonal Decomposition

Vincent Heuveline
(joint work with M. Hinze (TU Dresden))

We present an effective adjoint-based a-posteriori goal-oriented error control mechanism [1] for time integration of partial differential equations. The sensitivity information is obtained from the adjoint of a reduced order model of the full partial differential equations [2] while the reduced order model is adapted during the error estimation process. Several numerical examples illustrate the performance of the method.

References

[1] R. Becker, R. Rannacher, An optimal control approach to a posteriori error estimation in finite element methods, In Acta Numerica 2001, Cambridge University Press, pp. 1–101.

[2] M. Hinze, K. Kunisch, Three control methods for time-dependent fluid flow, Flow, Turbulence and Combustion 65, 273-298 (2000).

Model Reduction (and Spacetime Multigrid) in Control of Time-dependent PDEs

Michael Hinze

(joint work with K. Afanasiev (ZIB Berlin) and G. Büttner (TU Berlin))

We present an effective control method for mathematical models governed by systems of nonlinear time-dependent partial differential equations. It takes account of the fact that control inputs may alter the regime of the underlying physical process. The method in an adaptive manner constructs a hierarchy of appropriate low dimensional approximations to the mathematical model which then are used as subsidiary condition in the related optimization problem. We discuss different possibilities to construct low dimensional systems and the related modes (eigenfunctions of stationary problem, eigenfunctions of the linearized model and snapshot form of proper orthogonal decomposition).

As numerical example we present tracking-type control of the incompressible Navier-Stokes system as mathematical model for periodic flow around a cylinder. The numerical results of the adaptive approaches for different modes are compared. Furthermore they are compared to the result of the optimal control approach applied to the full Navier-Stokes system. It turns out that the quality of the controls obtained from the suboptimal approaches compares to that obtained by optimal control, and the computational costs for the optimal approach are at least one order of magnitude larger. More specifically, for the numerical example considered we obtain

$$\text{Runtime(Optimization)} = \text{(6-8) x Runtime(pdesolve)},$$

see [1].

In the second part of the talk we present preliminary results for multigrid in spacetime following the integral equation approach of Hackbusch. It turns out that for the numerical solution of linear-quadratic control problems we achieve

$$\text{Runtime(Optimization)} = \text{(7-8) x Runtime(pdesolve)},$$

for the numerical solution of a nonlinear problem with inexact Newton methods and spacetime multigrid for the Newton defect system we achieve

$$\text{Runtime(Optimization)} = \text{20 x Runtime(pdesolve)},$$

see [2].

References

[1] K. Afanasiev, M. Hinze, Adaptive control of a wake flow using proper orthogonal decomposition, Lecture Notes in Pure and Applied Mathematics 216 (2001).

[2] G. Büttner, Ein Mehrgitterverfahren zur optimalen Steuerung parabolischer Probleme, Dissertation TU Berlin (2004).

Adaptive Finite Elements for Inverse Scattering

Claes Johnson
(joint work with L. Beilina)

We apply an adaptive hybrid FEM/FDM method for an inverse scattering problem for the time-dependent acoustic wave equation in 2D and 3D where we seek to reconstruct an unknown sound velocity $c(x)$ from measured wave-reflection data. Typically this corresponds to identifying an unknown object [scatterer] in a surrounding homogeneous medium.

We use an optimal control approach where we seek a sound velocity $c[x]$ which minimizes the difference between computed and measured output data in a discrete L_2 norm. We solve the optimization problem by a quasi-Newton method where in each step we compute the gradient by solving a forward [direct] and an backward [adjoint] wave propagation problem.

To compute the backward and forward wave propagation problems we use an adaptive hybrid finite element/finite difference method where we exploit the flexibility of mesh refinement and adaption of the finite element method in a domain covering the object and the efficiency of a structured mesh finite difference method in the surrounding homogeneous domain. The hybrid scheme can be viewed as a finite element scheme on a partially unstructured mesh which gives a stable coupling of the two methods.

We use an adaptive mesh refinement algorithm to improve the accuracy of the reconstruction and speed up the convergence of the quasi-Newton method.

References

[1] L. Beilina, C. Johnson, Adaptive hybrid FEM/FDM methods for inverse scattering problems, Preprint, Chalmers Finite Element Center, 2003.

Structural Aspects for Numerical Methods in Optimal Control of Evolution Equations

Karl Kunisch

Motivated by optimal control problems in fluid dynamics describing vortex or drag reduction I remark in this survey talk on some aspects in open as well as closed loop numerical optimal control.

In the first part the difference between the optimization based approach and methods focused on solving the optimality system is explained. Sequential quadratic programming methods are compared to Newtons method. In the former the linearized state equation in the latter the nonlinear state equations are solved, resulting in primal feasibility in case of Newtons method [HK2]. In view of time-stepping techniques which are typically used to integrate the dynamical system, the numerical cost between the solution of the linearized equation and inexact

solutions to the nonlinear equation can be low. This suggests to favor the Newton over the SQP method for optimal control of evolution problems. — Finally a new cost functional for vortex suppression relying on local phase plane analysis is proposed [SK].

The second part is devoted to feedback control. For linear quadratic problems feedback control is completely characterized by means of an operator Riccati equation. In the nonlinear case the feedback control relies on the viscosity solution of a Hamilton-Jacobi-Bellman equation, which is numerically unfeasible unless the dimension of the (discretized) state-space is unreasonably small. For optimal control of fluid flow this requires the use of approximation strategies (beyond state space discretization). Here we explain techniques which rely on model reduction based on proper orthogonal decomposition (POD) combined with numerical solutions of the HJB-equation for infinite as well as finite horizon problems for the reduced problem [KV1, KV2, KX]. We also describe receding horizon techniques which rely on a time domain splitting strategy. To analytically justify their use we consider the stabilization problem of steady states. If a control Ljapunov functional is used as terminal penalty then the receding horizon control successfully drives the trajectory to the steady state [IK1, IK2].

References

[HK1] M. Hinze, K. Kunisch, Three control methods for time-dependent fluid flow, Journal of Flow, Control and Combustion, 65 (2000), 273–298.

[HK2] M. Hinze, K. Kunisch, Second order methods for optimal control of time-dependent fluid flow, SIAM J. Control and Optimization, 40 (2001), 925–946.

[IK1] K. Ito, K. Kunisch, On asymptotic properties of receding horizon optimal control, SIAM J. on Control and Optimization, 40 (2001), 1455–1472.

[IK2] K. Ito, K. Kunisch, Receding horizon control for infinite dimensional systems, ESAIM, Control, Optimization and Calculus of Variations, 8 (2002), 741-760.

[IK3] K. Ito, K. Kunisch, Semi-smooth Newton methods for state-constrained optimal control problems, to appear in Systems and Control Letters.

[KM] K. Kunisch, X. Marduel, Sub-optimal control of transient non-isothermal viscoelastic fluid flows, Physics of Fluids, 13(2001), 2478–2491.

[KV1] K. Kunisch, S. Volkwein, Galerkin proper orthogonal decomposition methods for parabolic systems, Numerische Mathematik, 90 (2001), 117–148.

[KV2] K. Kunisch, S. Volkwein, Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics, to appear in SIAM J. on Numerical Analysis.

[KX] K. Kunisch, L. Xie, Suboptimal feeback control of flow separation by POD Model Reduction, Submitted.

[SK] Y. Spasov, K. Kunisch, Dynamical system based optimal control of incompressible fluids: Boundary control, Submitted.

Adaptive Wavelets for Optimal Control Problems

Angela Kunoth

For the fast numerical solution of control problems governed by partial differential equations, an adaptive algorithm based on wavelets is proposed. The framework allows for linear elliptic and parabolic PDEs in full weak space–time

formulation as constraints as well as for problems with distributed or with Neumann or Dirichlet boundary control.

A quadratic cost functional which may involve fractional Sobolev norms of the state and the control is to be minimized subject to linear constraints in weak form. Placing the problem into the framework of (biorthogonal) wavelets allows us to represent the functional and the constraints in terms of ℓ_2–norms of wavelet expansion coefficients and constraints in form of an ℓ_2 automorphism. The resulting first order necessary conditions are then derived as a (still infinite) system in ℓ_2.

Applying the machinery developed in [CDD], we propose an adaptive method for the coupled system for the state, adjoint and control variables. An essential ingredient is that the scheme can be interpreted as an inexact gradient descent method, where in each iteration step the primal and the adjoint system needs to be solved up to a prescribed accuracy. In particular, the method resolves each of the involved variables separately, without having to resort to a common underlying grid.

The convergence analysis of the algorithm is crucially based on the fact that the wavelet framework allows us to step by step break down an ideal iteration on the infinite system to computable quantities. Thus, the method captures all relevant features from the original control problem with respect to infinite function spaces and specifically resolves any type of singularity coming from the data or the domain. Consequently, the approximate solutions generated by the adaptive algorithm can be shown to converge (in the energy norm) to the exact solution triple (state, adjoint state, control) for any prescribed accuracy.

Moreover, it is shown that the adaptive algorithm is asymptotically optimal, that is, the convergence rate achieved for computing the solution up to a desired target tolerance is asymptotically the same as the wavelet–best N–term approximation of the solution, and the total computational work is proportional to the number of computational unknowns.

References

[CDD] A. Cohen, W. Dahmen, R. DeVore, Adaptive wavelet methods for elliptic operator equations — Convergence rates, Math. Comp., 70 (2001), 27–75.

[DK] W. Dahmen, A. Kunoth, Adaptive wavelet methods for linear–quadratic elliptic control problems: Convergence Rates, Preprint #46, SFB 611, Universität Bonn, December 2002, submitted for publication, revised, November 2003.

[GK] M. Gunzburger, A. Kunoth, Adaptive wavelets for optimal control of linear parabolic PDE's, in preparation.

[K] A. Kunoth, Adaptive wavelet schemes for an elliptic control problem with Dirichlet boundary control, Preprint #109, SFB 611, Universität Bonn, November 2003, to appear in Numer. Algor.

A Paradigm for Adaptivity and Optimal Control

Rolf Rannacher

We present a general approach to "goal-oriented" error estimation and mesh adaptation in the context of optimal control problems of the abstract form

$$J(u,q) = \min!, \qquad A(u,q)(\cdot) = 0,$$

where the equation constraint is a PDE. The starting point is the first-order optimality condition, the so-called KKT system,

$$\nabla_{u,q,z} L(u,q,z) = 0, \qquad L(u,q,z) := J(u,q) - A(u,q)(z),$$

obtained by the classical Euler-Lagrange approach. This set of nonlinear and strongly coupled PDEs for the primal variable (state) u, the control variable q (control), and the dual variable z (adjoint) is approximated by a Galerkin finite element method. The topic of this talk is a strategy for the a posterori construction of finite element meshes which are most economical for the optimization process. Exploiting the particular structure of the KKT system an error representation is derived for the discretization with respect to the cost functional in terms of the residuals of the computed solution and a remainder term which is cubic in the approximation errors,

$$J(u,q) - J(u_h,q_h) = \tfrac{1}{2}\nabla_{u,q,z} L(u_h,q_h,z_h)(z-\psi_h, q_h-\chi_h, u-\phi_h) + R_h^{(3)}(e^u,e^q,e^z).$$

The remainder term is usually neglected. The computational evaluation of the residual term yields "weighted" a posteriori error estimates of the form

$$J(u,q) - J(u_h,q_h) \approx \sum\nolimits_{K\in T_h} \left\{ \rho_K^u \omega_K^z + \rho_K^q \omega_K^q + \rho_h^z \omega_K^u \right\},$$

which can guide the mesh adaptation process. In these estimates "primal" residuals and "dual" weights as well as "dual" residuals and "primal" weights are crosswise multiplied, while in the traditional approach these quantities are added. Therefore, this approach of mesh adaptation is called "Dual Weighted Residual (DWR) Method". The resulting meshes for discretizing the KKT system are most economical since only that information is represented which is really needed for the optimization process. When an optimal control $q_h^{\rm opt}$ is determined for this sparse model, using for instance a Newton-type iteration, a more accurate primal solution $\tilde{u}_h^{\rm opt}$ can be generated in a post-processing step by solving the state equation once again on a finer mesh. The performance of the DWR method is demonstrated for a simple diffusion-reaction problem with Neumann-boundary control. The very same approach can also be used in the context of time discretization and eigenvalue problems, e.g., in the control of the Navier-Stokes equations for drag minimization and for the stability analysis of the resulting optimal states. For further information and other classes of applications, we refer to the literature listed below.

References

[1] R. Becker, H. Kapp, R. Rannacher, Adaptive finite element methods for optimal control of partial differential equations: Basic concepts, SIAM J. Control Optim. 39, 113–132 (2000).
[2] R. Becker, R. Rannacher, An optimal control approach to error estimation and mesh adaptation in finite element methods, Acta Numerica 2001, pp. 1–101, Cambridge University Press.
[3] R. Becker, Mesh adaptation for stationary flow control, J. Math. Fluid Mech. 3, 317–341 (2001).
[4] R. Becker, V. Heuveline, R. Rannacher, An optimal control approach to adaptivity in computational fluid mechanics, Int. J. Numer. Meth. Fluids. 40, 105–120 (2002).
[5] W. Bangerth, R. Rannacher, Adaptive Finite Element Methods for Differential Equations, Lectures in Mathematics, ETH Zürich, Birkhäuser, Basel-Boston-Berlin, 2003.

Modelling and Globalizing Applied to Optimal Control Algorithms

Ekkehard W. Sachs

Various models in algorithms for optimization have led to different strategies of globalizing optimization methods. The trust region concept is a general globalization strategy which can be applied to various model functions.

For an unconstrained optimization problem the trust region strategy defines at each iteration point u_k a nonlinear model function $m_k(u_k + \cdot)$ which is minimized over a ball of radius Δ. The radius is updated according to the accuracy of the value of the model function in comparison to the value of the true function at the candidate at the next iteration point. The trust region is updated in a similar fashion. These strategies together with a condition on a sufficient descent for a function value yield a global convergence statement.

The linear model leads to the well known Armijo step size rule, whereas the quadratic model yields a smooth transition from steepest descent to the Newton step. It is less known that the trust region method can also be applied to nonlinear models as pointed out by Toint [4].

The training of neural networks is one area where these techniques have been applied successfully for large scale problems.

Another application is reduced order modelling, in particular, proper orthogonal decomposition for the optimal control of Navier-Stokes equation. Afanasiev and Hinze [1] also use adaptive techniques to adjust the POD model during the iteration. In Arian, Fahl, Sachs [2, 3] the trust region approach is utilized to control the adaption of the POD model. In this case the differential equation is replaced by a reduced order model of smaller size. In this case the original function $f(y(\cdot))$ is replaced by the model function $m_k(u_k + \cdot) = f(y_{POD,u_k}(\cdot))$.

Numerical examples are given for an example of the control of a driven cavity flow problem.

References

[1] K. Afanasiev, M. Hinze, Adaptive control of a wake flow using proper orthogonal decomposition, Lecture Notes in Pure and Applied Mathematics 216 (2001).

[2] E. Arian, M. Fahl, and E.W. Sachs, Trust-region proper orthogonal decomposition for flow control, ICASE Report 2000-25, ICASE, NASA Langley Research Center, Hampton, 2000.
[3] E. Arian, M. Fahl and E.W. Sachs, Managing POD Models by optimization methods, IEEE CDC Conference Proceedings, Las Vegas, 2002.
[4] P.L. Toint, Global convergence of a class of trust region methods for nonconvex minimization in Hilbert space, IMA J. Num. Anal., 8 (2) 231-252 (1988).

ODE Concepts for PDE Optimization

Volker H. Schulz
(joint work with S. Hazra (U Trier))

In this talk, we exploit algorithmic concepts from the ODE world for the solution of optimization problems with PDE constraints. In particular, methodological results for two specific application problems are reported.

In the first application problem, we solve parameter identification problems for instationary multiphase flow in the subsurface. Multiple shooting in combination with a reduced Gauss-Newton approach, due to Bock and Schloeder (1981-), yields an efficient and robust algorithm with low storage requirements (cf. [1] and several subsequent papers).

In the second application problem, we study shape optimization for the design of parts of the surface of airplanes under drag optimization. First results regarding a new collaborative project together with DLR Braunschweig, Airbus Germany, EADS and others, which has started recently, are reported. The algorithmic workhorse is a generalization of reduced SQP techniques to continuous reduced SQP techniques in a pseudo-timestepping framework.

Both applications show that one can profit from knowledge of ODE concepts for optimization problems also in a PDE context.

References

[1] S. Hazra, V. Schulz, Numerical identification in multiphase flow through porous media, Comput. Vis. Sci. 5:107-113 (2002).
[2] S. Hazra, V. Schulz, Simultaneous pseudo-timestepping for PDE-model based optimization problems, Report 02-23, U Trier, 2002, submitted to BIT.
[3] S. Hazra, V. Schulz, Continuous reduced SQP methods for airwing design, in preparation.

On Optimal Control Problems with Pointwise State Constraints

Fredi Tröltzsch

We consider the parabolic optimal control problem with pointwise constraints on the control and the state,

$$\min J(y,u) = \frac{1}{2}\,\|y(T) - y_d\|^2_{L^2(\omega)} + \frac{\nu}{2}\,\|u\|^2_{L^2(Q)}$$

subject to

$$y_t - \Delta y + d(y) = u \qquad \text{in } Q = \Omega \times (0,T)$$

with homogeneous Neumann boundary condition and initial condition $y(0) = y_0$. Moreover, pointwise constraints are imposed on the control $u \in L^\infty(Q)$ and the state $y \in Y := W(0,T) \cap C(\bar{Q})$,

$$\begin{aligned} 0 &\le u(x,t) \le b(x,t) \\ 0 &\le c(x,t) + \gamma\, y(x,t), \end{aligned}$$

$(x,t) \in Q$. The following data are given: A bounded domain $\Omega \subset \mathbb{R}^N$ with $C^{1,1}$-boundary Γ, functions $y_d,\, y_0 \in C(\bar{\Omega})$, $b,\, c \in C(\bar{Q}), d \in C^{2,1}(\mathbb{R})$ with $d'(y) \ge 0$, and $\gamma > 0$, $\nu > 0$. For all $u \in L^\infty(Q)$, a unique state $y = y(u)$ exists in Y. Let $\bar{u}$ be locally optimal in the sense of $L^\infty(Q)$ and define $\bar{y} = y(\bar{u})$. Then the associated first order necessary optimality conditions can be formulated on using the Lagrange functional

$$\mathcal{L}(y,u,p,\mu) = J(y,u) - \int_Q [(y_t + d(y) - u)p + \nabla y \cdot \nabla p]\, dxdt - \int_Q (y+c)\, d\mu(x,t),$$

where p stands for the adjoint state and $\mu \in C(\bar{Q})^\star$ is the Lagrange multiplier associated with the state constraints. If a constraint-qualification is satisfied at $(\bar{y}, \bar{u})$, then a non-negative Borel measure μ exists such that the first order necessary optimality conditions

$$\begin{aligned} \tfrac{\partial \mathcal{L}}{\partial y}(\bar{y},\bar{u},p,\mu)\, y &= 0 \quad \forall\, y \in W(0,T) \text{ with } y(0) = 0 \\ \tfrac{\partial \mathcal{L}}{\partial u}(\bar{y},\bar{u},p,\mu)(u-\bar{u}) &\ge 0 \quad \forall u \in L^\infty(\Omega) \text{ with } 0 \le u \le b \\ \int_Q (y+c)\, d\mu &= 0 \end{aligned}$$

are satisfied. This follows from results [2] and [5]. Assume conversely that a pair $(\bar{y}, \bar{u})$ is given that satisfies all constraints and the first-order necessary conditions. One might expect that the following standard condition is sufficient for local optimality of $\bar{u}$:

(SSC) There is $\delta > 0$ such that

$$\mathcal{L}''(\bar{y},\bar{u},p,\mu)(y,u)^2 = \|y(T)\|^2_{L^2(\Omega)} + \nu \|u\|^2_{L^2(Q)} - \int_Q d''(\bar{y}) p\, y^2\, dxdt \ge \delta\, \|u\|^2_{L^2(Q)}$$

holds for all pairs $(y,u) \in Y \times L^\infty(Q)$ satisfying the state equation linearized at $(\bar{y}, \bar{u})$.

Then $\bar{u}$ is expected to be locally optimal with respect to the L^2-topology. This holds true, if the mapping $u \mapsto y(u)$ is continuous from L^2 to $C(\bar{Q})$, i.e. for $N = 1$ in our parabolic example with distributed control. For the case of Neumann boundary control, this is not true.

If, however, the state-constraints are deleted, then (SSC) is sufficient for local optimality in the L^∞-topology for all N.

Unfortunately, we have not been able to prove the same result with pointwise state constraints. Local optimality can still be shown in the sense of L^∞ for $N = 2$. For $N > 2$ we cannot prove that (SSC) is really sufficient for local optimality. The obstacle is the need to estimate $\mathcal{L}''(\bar{y},\bar{u},p,\mu)(y,u)^2$ against the L^2-norm of u. Due

to the presence of the measure μ in the right-hand side of the adjoint equation, p is not bounded in the case of state-constraints. Therefore, the estimation of the third quantity in the expression for $\mathcal{L}''$ causes troubles.

The situation is even worse, if (SSC) is weakened by considering also strongly active state-constraints. Then only for $N = 1$ the local optimality can be shown in the case of distributed control, while boundary controls cannot be handled at all. We refer to [6]. The results are slightly better for elliptic problems. L^2-optimality can be shown for $N \leq 3$ and distributed control (since $H^2(\Omega) \subset C(\bar{\Omega})$) and $N = 2$ for Neumann boundary control ($H^{3/2-\varepsilon}(\Omega) \subset C(\bar{\Omega})$), [3]. To overcome these difficulties, a Lavrentiev type regularization is suggested. Consider the mixed-pointwise control-state constraints

$$-\varepsilon\, u(x,t) \leq c(x,t) + \gamma\, y(x,t).$$

In this case, the existence of an associated regular Lagrange multiplier $\mu \in L^\infty(Q)$ can be shown, [1]. Moreover, this concept is useful for numerical approximations. For $\varepsilon \downarrow 0$, the associated optimal control converges to $\bar{u}$. This is demonstrated for linear-quadratic elliptic problems with $N = 2$, [4].

References

[1] M. Bergounioux, F. Tröltzsch, Optimal control of semilinear parabolic equations with state-constraints of bottleneck type, ESAIM, Control, Optimisation and Calculus of Variations, 4 (1999), 595–608.

[2] E. Casas, Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations, SIAM J. Control Optimization 35 (1997), 1297–1327.

[3] E. Casas, F. Tröltzsch, and A. Unger, Second order sufficient optimality conditions for some state-constrained control problems of semilinear elliptic equations, SIAM J. Control and Optimization, 38 (2000), 1369–1391.

[4] C. Meyer, A. Rösch, and F. Tröltzsch, Optimal control problems of PDEs with regularized pointwise state constraints, Tech. Report 14-2003, Inst. of Math., TU Berlin. Submitted.

[5] J.-P. Raymond, H. Zidani, Hamiltonian Pontryagin's principles for control problems governed by semilinear parabolic equations, Applied Mathematics and Optimization 39 (1999), 143–177.

[6] J.-P. Raymond and F. Tröltzsch, Second order sufficient optimality conditions for nonlinear parabolic control problems with state constraints, Discrete and Continuous Dynamical Systems, 6 (2000), 431–450.

Optimal Control of Nonlinear Parabolic Systems by Second Order Methods

Stefan Volkwein

In the talk three different optimal control problems for nonlinear parabolic systems are considered.

The first example is concerned with optimal boundary control of an instationary reaction-diffusion system in three spatial dimensions. This problem involves a coupled nonlinear system of parabolic differential equations with bilateral as well as integral control constraints. We include the integral constraint in the cost by

a penalty term whereas the bilateral control constraints are handled explicitly. A primal-dual active set strategy is utilized to compute optimal solutions numerically. The algorithm is compared to a semi-smooth Newton method.

As a second example an optimal control boundary problem for the Stefan problem is considered. Here, an inexact Newton method is applied with quasi-Newton approximations for the Hessian. To ensure positivity of the Hessian, a line search based on the Wolfe-Powell conditions is utilized.

Finally, laser surface hardening of steel is formulated in terms of an optimal control problem, where the state equations are a semilinear heat equation and an ordinary differential equation, which describes the evolution of the high temperature phase. To avoid the melting of the steel we have to impose state constraints for the temperature. Including the state constraints into the cost functional by a penalty term, a globalized SQP method with a reduced Hessian is applied to solve the control problem numerically. To ensure the convergence of the algorithm a numerically inexpensive globalization strategy is used.

References

[1] R. Griesse, S. Volkwein, Analysis for optimal boundary control for a three-dimensional reaction-diffusion system, Technical Report No. 286, Special Research Center F 003 Optimization and Control, Project area Continuous Optimization and Control, University of Graz & Technical University of Graz, Dezember 2003.

[2] M. Hintermüller, K. Ito, and K. Kunisch, The primal-dual active set strategy as a semi-smooth Newton method, SIAM J. Optimization, 13:865-888, 2003.

[3] K. Ito, K. Kunisch, The primal-dual active set method for nonlinear problems with bilaterally constraints, Technical Report No. 214, Special Research Center F 003 Optimization and Control, Project area Continuous Optimization and Control, University of Graz & Technical University of Graz, 2001, to appear in SIAM J. Control and Optimization.

[4] M. Ulbrich, Semismooth Newton methods for operator equations in function spaces, SIAM J. Optimization, 13:805-842, 2003.

[5] S. Volkwein, A globalized SQP method for the optimal control of laser surface hardening, Technical Report No. 272, Special Research Center F 003 Optimization and Control, Project area Continuous Optimization and Control, University of Graz & Technical University of Graz, June 2003, submitted.

Interior–Point Methods for Optimization Problems with PDE Constraints

Beate Winkelmann

(joint work with R.E. Bank and Ph.E. Gill)

Methods are proposed for the numerical solution of optimal control problems with partial differential equation (PDE) constraints and inequality constraints on

the control variable. The general form of the problem is:

$$\underset{y,u}{\text{minimize}}\ \rho(y,u) = \int_\Omega p_1(y,\nabla y,u)\,dx + \int_{\partial\Omega} p_2(y,\nabla y,u)\,ds$$

$$\text{subject to } \langle E(y,u),v\rangle = 0, \quad \forall v \in H,$$
$$\underline{u} \le u(x) \le \overline{u},$$

where $\Omega \subset \mathbb{R}^d$, $d \in \{1,2\}$, y is the state variable, u is the control variable, H is an appropriate function space, and $\langle E(y,u),v\rangle = 0$ is the weak form of a partial differential equation in divergence form. State and control variables are discretized using an adaptive finite-element approach. Algorithms for optimization and PDEs are combined to solve a discretized optimization problem over a sequence of adaptive meshes.

An interior-point method is used for the optimization part of the algorithm. The two main types of interior-point method are *primal methods* and *primal-dual methods*. As the names suggest, primal methods iterate over the primal variables only, while primal-dual methods iterate over the primal and dual variables simultaneously. However, both methods approach the solution by following a continuous path that approaches the solution from the interior of the set of admissible solutions. For a primal method the path is the trajectory of solutions of a sequence of equality-constrained problems parameterized by a scalar μ. Primal-dual methods define the path as the trajectory of points satisfying the perturbed first-order optimality conditions for the constrained problem.

Primal-dual methods are usually preferred for general constrained optimization because of their rapid convergence near the solution. However, in the PDE context, the use of a primal-dual method requires the adaptive discretization of both the primal and dual variables, which can lead to serious numerical difficulties if the dual variables are not sufficiently smooth. It is shown that these difficulties may be avoided if the primal method is implemented using an extrapolation scheme for the parameter μ.

Regardless of the particular choice of interior-point method, the linear systems to be solved at each iteration are large, symmetric and have PDE-like structure. These systems also become increasingly ill-conditioned as the solution is approached. In order to handle the size and sparsity of these systems, a preconditioned Krylov-space method is used. The choice of preconditioner is crucial for the performance of the optimization since the cost of solving the linear system dominates the overall cost of the computation. A good preconditioner lowers the cost of solving the linear systems significantly. The aim is to incorporate as much information as possible into the preconditioner without dramatically increasing the cost of the computation. To this end, the preconditioner has the same block structure as the original system and existing multilevel PDE preconditioners are used for some of the blocks. In particular, an algebraic multigrid preconditioner with

ILU smoothing is used for the PDE constraint blocks, and a symmetric Gauss-Seidel preconditioner is used for the control block and the full linear system. The preconditioner is fully parallel.

The PDE part of the algorithm uses adaptive mesh refinement based on an *a posteriori* hierarchical basis error estimator for the state variables. The path-following parameter μ and the PDE parameters are chosen to allow the discretization error and optimization error to go to zero at the same rate. An error estimator based on the state variable allows the mesh to be adaptively refined and unrefined without the additional cost of solving the adjoint equation.

These ideas are illustrated in the context of the elliptic finite-element PDE package PLTMG. Numerical results are presented for a particular optimal control problem involving an elliptic PDE constraint. The adaptive refinement algorithm requires the solution of an optimization problem with up to 22.5M variables. This problem was solved on a 256 processor Beowulf cluster in approximately 8 minutes.

References

[1] R.E. Bank, M. Holst, A new paradigm for parallel adaptive mesh refinement, SIAM J. Sci. and Statist. Comput., 22(4):1411–1443, 2000.

[2] R.E. Bank, PLTMG: A Software Package for Solving Elliptic Partial Differential Equations, Users' Guide 8.0, Software, Environments and Tools, Vol. 5, SIAM, Philadelphia, 1998.

[3] R. E. Bank, P. E. Gill, R. F. Marcia, Interior methods for a class of elliptic variational inequalities, In L.T. Biegler et al. (eds.), Large-scale PDE-constrained Optimization, vol. 30, Lecture Notes in Computational Science and Engineering, 218–235, Berlin, Heidelberg and New York, 2003. Springer-Verlag.

[4] A. Forsgren, P.E. Gill, M.H. Wright, Interior methods for nonlinear optimization, SIAM Rev., 44:525–597, 2002.

[5] B. Winkelmann, Interior-point Methods and their Application to Optimal Control with PDE Constraints, University of California San Diego, 2004, Ph.D. Thesis, in preparation.

Reporter: Angela Kunoth

Participants

Dr. Roland Becker
roland.becker@univ-pau.fr
roland.becker@iwr.uni-heidelberg.de
Laboratoire de Mathematiques
Appliquees
Universite de Pau et des Pays de l'Adour
BP 1155
F-64013 Pau Cedex

Dr. George Biros
biros@seas.upenn.edu
Dept. of Computer and Information
Science, University of Pennsylvania
220 Towne Building
220 S. 33rd. St.
Philadelphia PA 19104-6315 – USA

Prof. Dr. Hans Georg Bock
scicom@iwr.uni-heidelberg.de
Bock@IWR.Uni-Heidelberg.De
Interdisziplinäres Zentrum
für Wissenschaftliches Rechnen
Universität Heidelberg
Im Neuenheimer Feld 368
D–69120 Heidelberg

Dr. Alfio Borzi
alfio.borzi@uni-graz.at
Institut für Mathematik
Karl-Franzens-Universität Graz
Heinrichstr. 36
A-8010 Graz

Carsten Burstedde
burstedde@iam.uni-bonn.de
Institut für Angewandte Mathematik
Universität Bonn
Wegelerstr. 6
D–53115 Bonn

Prof. Omar Ghattas
oghattas.@cs.cmu.edu
Carnegie Mellon University
Department of Civil and
Environmental Engineering
Pittsburgh PA 15213-3890 – USA

Prof. Dr. Andreas Griewank
griewank@math.hu-berlin.de
Institut für Mathematik
Humboldt-Universität Berlin
Unter den Linden 6
D–10099 Berlin

Prof. Dr. Max D. Gunzburger
gunzburg@csit.fsu.edu
CSIT/DSL 400
School of Computational Science and
Information Technology
Florida State University
Tallahassee FL 32306-4120 – USA

Dr. Vincent Heuveline
vincent.heuveline@iwr.uni-heidelberg.de
Institut für Angewandte Mathematik
Universität Heidelberg
Im Neuenheimer Feld 294
D–69120 Heidelberg

Prof. Dr. Michael Hinze
hinze@math.tu-dresden.de
TU Dresden
Institut für Numerische Mathematik
Willersbau C 312
D–01062 Dresden

Prof. Dr. Claes Johnson
claes@math.chalmers.se
Department of Mathematics
Chalmers University of Technology
S-412 96 Göteborg

Prof. Dr. Karl Kunisch
karl.kunisch@uni-graz.at
Institut für Mathematik
Karl-Franzens-Universität Graz
Heinrichstr. 36
A-8010 Graz

Prof. Dr. Angela Kunoth
kunoth@iam.uni-bonn.de
Institut für Angewandte Mathematik
Universität Bonn
Wegelerstr. 6
53115 Bonn

Prof. Dr. Günter Leugering
leugering@am.uni-erlangen.de
Lehrst. f. Angewandte Mathematik II
Universität Erlangen-Nürnberg
Martensstr. 3
D–91058 Erlangen

Daniel Oeltz
oeltz@iam.uni-bonn.de
Institut für Angewandte Mathematik
Universität Bonn
Wegelerstr. 6
D–53115 Bonn

Prof. Dr. Rolf Rannacher
rannacher@iwr.uni-heidelberg.de
Institut für Angewandte Mathematik
Universität Heidelberg
Im Neuenheimer Feld 294
D–69120 Heidelberg

Prof. Dr. Ekkehard Sachs
sachs@uni-trier.de
Abteilung Mathematik
Fachbereich IV
Universität Trier
D–54286 Trier

Prof. Dr. Volker Schulz
Volker.Schulz@uni-trier.de
Fachbereich IV - Mathematik
Numerik - Optimierung, partielle
Differentialgleichungen
Universität Trier
D–54286 Trier

Prof. Dr. Fredi Tröltzsch
troeltzsch@math.tu-berlin.de
Institut für Mathematik
Technische Universität Berlin
Sekr. MA 4-5
Strasse des 17. Juni 136
D–10623 Berlin

Doz.Dr. Stefan Volkwein
stefan.volkwein@uni-graz.at
Institut für Mathematik
Karl-Franzens-Universität Graz
Heinrichstr. 36
A-8010 Graz

Beate Winkelmann
bmwinkelmann@ucsd.edu
Dept. of Mathematics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0112 – USA