Numerical Optimization of Multifunctional Components

Anthony T. Patera

Massachusetts Institute of Technology
Dept. of Mechanical Engineering
Room 3-266
77 Massachusetts Avenue
Cambridge, MA 02139

Air Force Office of Scientific Research
AFOSR/NM
801 North Randolph, Room 732
Arlington, VA 22203-1977

Approved for public release;
Distribution unlimited.

Engineering components are increasingly complex in composition and structure and increasingly multifunctional: indeed, it is only through complexity and multifunctionality that we can satisfy the stringent performance requirements associated with critical defense applications. However, these complex, multifunctional systems no longer admit intuitive analysis of trade-off considerations: we must pursue optimization — optimal choice of material, configuration, and deployment — to realize the potential of these new approaches. The essential mathematical enabler — and our focus in this project — is very fast yet reliable prediction of component behavior; armed with the latter, we may then pursue extensive optimization and even real-time adaptive design and control.

Critical ingredients of our approach are: (i) reduced-basis approximations to effect significant reduction in state-space dimensionality; (ii) \textit{a priori} error bounds to provide rigorous error estimation and control; (iii) “offline/online” computational decompositions to permit rapid evaluation of output bounds \textit{in the limit of many queries}. In this project, we extend our basic methodology to non-coercive, non-affine, non-linear, and “non-elliptic” (parabolic) problems; we may thus now address the full range of disciplines that typically describe actual thermo-structure-fluid-acoustic-electromagnetic multifunctional components.
Numerical Optimization of Multifunctional Components

Grant Number F49620-03-1-0356

Anthony T. Patera
Department of Mechanical Engineering
Massachusetts Institute of Technology

Background

Engineering components are increasingly complex in composition and structure and increasingly multifunctional: indeed, it is only through complexity and multifunctionality that we can satisfy the stringent performance requirements associated with critical defense applications. However, these complex, multifunctional systems no longer admit intuitive analysis of trade-off considerations: we must pursue optimization — optimal choice of material, configuration, and deployment — to realize the potential of these new approaches.

The optimization, control, and characterization of an engineering component or system requires the prediction of certain “quantities of interest,” or performance metrics, which we shall denote outputs — for example deflections, maximum stresses, maximum temperatures, heat transfer rates, flow rates, or lifts and drags. These outputs are typically expressed as functionals of field variables associated with a parametrized partial differential equation which describes the physical behavior of the component or system. The parameters, which we shall denote inputs, serve to identify a particular “configuration” of the component: these inputs may represent design or decision variables, such as geometry — for example in optimization studies; control variables, such as actuator power — for example in real-time applications; or characterization variables, such as physical properties — for example in inverse problems. We thus arrive at an implicit input-output relationship, the evaluation of which demands solution of the underlying partial differential equation.

Our goal is the development of computational methods that permit rapid and reliable evaluation of this partial-differential-equation-induced input-output relationship in the limit of many queries — that is, in the design, optimization, control, and characterization contexts. The “many query” limit has certainly received considerable attention. Our approach is based on the reduced-basis method, first introduced in the late 1970s for nonlinear structural analysis [1,2], and subsequently developed more broadly in the 1980s and 1990s [3–7]. The reduced-basis method recognizes that the field variable is not, in fact, some arbitrary member of the infinite-dimensional solution space associated with the partial differential equation; rather, it resides on a much lower-dimensional manifold induced by the parametric dependence.

The reduced-basis approach, as earlier articulated, is local in parameter space in both practice and theory. To wit, Lagrangian or Taylor approximation spaces for the low-dimensional manifold are typically defined relative to a particular parameter point: the associated a priori
convergence theory relies on asymptotic (implicit function theorem) arguments in sufficiently small neighborhoods; the computational improvements — relative to conventional (say) finite element approximation — are often quite modest [4]. Our work [8–12] differs from these earlier efforts in several important ways: first, we develop (in some cases, provably) global approximation spaces; second, we introduce rigorous a posteriori error estimators; and third, we exploit offline/online computational decompositions (see [7] for an earlier application of this strategy within the reduced-basis context). These three ingredients allow us to reliably decouple the generation and projection stages of reduced-basis approximation, thereby effecting computational economies of several orders of magnitude.

First, the error in the reduced-basis approximation typically vanishes exponentially as a function of $N$, the dimension of the reduced-basis space: sufficient accuracy can thus be obtained with only $N = O(10) – O(100)$ degrees of freedom. Second, we can rigorously and sharply bound (a posteriori) the error in the reduced-basis approximation of the outputs of interest, thus permitting optimal truncation — selection of (close to) the smallest $N$ for which the desired error tolerance can be achieved. Third, we can decompose the computational effort into two stages: an expensive (offline) stage performed once; and an inexpensive (online) stage performed many times. The operation count for the online stage — in which, given a new value of the input, we calculate the output and associated error bound — depends only on $N$ (typically very small) and the parametric complexity of the operator. This very low marginal cost is critical in the optimization context.

**Particular Achievements**

*Non-coercive Problems*

In our earlier work, we treated primarily coercive operators; in particular, our earlier approaches did not permit rigorous a posteriori error estimation for non-coercive operators. We recently have developed [11] and subsequently refined [13; 14] procedures that now permit rigorous a posteriori error estimation for noncoercive problems — such as the Helmholtz equation (reduced-wave equation) that arises in acoustics, elastodynamics, and electromagnetics. The critical new ingredient is a lower bound for the inf-sup constant (singular value) that characterizes the stability of the operator — and hence controls the relationship between the dual norm of the residual and the error. Our (most efficient) approach is based on a piecewise-constant or piecewise-linear construction that exploits classical concavity properties of the inf-sup eigenproblem [14]. Examples of Helmholtz reduced-basis approximation and associated a posteriori error estimation — and application to inverse problems in elasticity — are given in [14].

*Non-linear Problems*

In our earlier work, we treated primarily linear operators; in particular, our earlier approaches [9] did not permit rigorous a posteriori error estimation for nonlinear problems (except in certain monotonic cases [11]). We have recently developed approaches — at present restricted to quadratic nonlinearities such as the incompressible Navier-Stokes equations describing fluid
flow (though see below) — that do indeed permit rigorous a posteriori output bounds. The
foundation is the Brezzi-Rappaz-Raviart theory for analysis of variational approximations of
nonlinear partial differential equations [15]. The key computational ingredients are efficient
offline-online evaluation of the dual norm of the residual (similar to the procedure we developed
earlier for linear problems); adaptation of the non-coercive inf-sup lower bound described above
to the nonlinear case; and quantitative calculation of the relevant Sobolev embedding constants
characterizing the nonlinearity. The method is described in detail in [14; 16], with application to
natural convection in an enclosure (a problem relevant to materials processing).

Non-affine Problems

Until recently, we could only treat partial differential equations that are (i) affine in the
parameter, and (ii) at most quadratically nonlinear in the state variable. Absent these
assumptions, our offline-online decomposition breaks down. In fact, both of these restrictions
can be addressed by a new “empirical interpolation” approach developed in collaboration with
Professor Yvon Maday of University Paris VI [17]. In this approach, we replace general non-
affine, nonlinear functions of the parameter, spatial coordinate, and state variable with collateral
reduced-basis expansions. The critical ingredients of the approach are (i) good collateral
reduced-basis samples and spaces, (ii) a stable and inexpensive online interpolation procedure by
which to determine the collateral reduced-basis coefficients (as a function of the parameter), and
(iii) effective a posteriori error bounds with which to quantify the effect of the newly introduced
truncation. With these techniques we can now efficiently and reliably consider many new
parameter variations (e.g., general shape optimization) as well as many new classes of nonlinear
phenomenon (e.g., radiation heat transfer [18]). (We note however that, in general, we do lose
some rigor in our a posteriori error bounds.)

Non-elliptic (Parabolic) Problems

We have recently extended our reduced-basis and a posteriori error estimation procedures to
parabolic problems — for example, the heat (transient conduction) equation. The essential
ingredient is the identification of a new (parameter, time) variable in which to develop the
reduced-basis approximation. The approach entails several key innovations relative to the
elliptic case both as regards approximation (e.g., impulse sampling procedures that permit
general temporal forcing functions — as arise in optimal control) and a posteriori error
estimation (adjoint techniques that provide bounds for the error in particular outputs of interest as
a function of time). The methods are further described — and applied to a model “blast
deflector” problem — in [14].

Adaptive Sampling

We have developed — in all the contexts above (including the temporal parabolic case) —
efficient adaptive sampling procedures that permit us to construct (greedily) optimal samples
that, in turn, yield rapidly convergent (and numerically stable) reduced-basis spaces. Given an
initial parameter sample and associated approximation space, we can readily calculate — thanks
to our fast error bounds — (an approximation to) both the maximum error over the parameter
space and the location of the largest errors; this then serves to select the next points to be
included in the sample (and, ultimately, to terminate the approximation upon satisfaction of a given “worst-case” tolerance). These methods are, in essence, “POD-like” economization procedures [19] in which we need never form the vast majority of the snapshots.

User Interfaces

In the past [20] we have developed an extensive Web repository system to house online servers and an associated suite of (thin) clients by which to access these codes. In addition, in the current project, we have developed a simpler variant of this idea: MATLAB .m files which — for a particular partial differential equation, output(s), and input domain — (self-)contain the databases and online codes that, at the usual MATLAB command-line level, can then be invoked to provide the desired output and output error bound for any given input. In essence, we reduce Navier-Stokes to the computational (and user) complexity — and certainty — of “sin(x).” (We note that the development of the MATLAB .m file is time consuming, and is furthermore specific to each particular application; however, once in place, the response to the “end user” is quite literally real-time.) Examples of these MATLAB .m files are of course available to the sponsor upon request.

Optimization

The most common approach to the optimization of systems described by partial differential equations is to combine state-of-the-art optimization techniques — such as pattern search techniques, Sequential Linear/Quadratic Programming (SLP/SQP) approaches, or Newton Interior Point Methods (IPM) — with state-of-the-art partial differential equation discretization techniques — such as the finite element method. The best approaches (e.g., [21]) consider the optimization formulation and partial differential equation treatment in an integrated fashion. However, even these “best methods” remain quite expensive — a sufficiently accurate discretization may require hundreds of thousands of degrees of freedom — with computational times often measured in hours or even days. Further exacerbating the situation, realistic design exercises typically require many optimization cycles — corresponding to variations in the design, operation, and environment parameters that define the objectives and constraints. In the multifunctional context, these difficulties are even further amplified by the presence of a variety of different equations, objectives, and constraints that rarely admit any overarching simple structure.

An alternative approach to the optimization of systems described by partial differential equation is to replace the (say) finite element discretization with a much lower order description. This description may take the form of a model based primarily on physical reasoning or empirical constructions, or of a formal approximation directly derived from the underlying partial differential equations. A variety of frameworks (e.g., [22]) have been proposed for the incorporation of these low order descriptions into the design and optimization context. The main outstanding difficulty is the lack of rigorous, sharp, inexpensive error estimators for the outputs which appear in the optimization statement. Absent such a measure of fidelity, either the model truncation must be chosen (arbitrarily) conservatively — thus compromising efficiency; or the model truncation may be chosen overly optimistically — thus compromising convergence, optimality, and feasibility (with respect to the exact mathematical description).
Our approach — reduced-basis-based optimization — is of course of the low-order-approximation variety. However, in contrast to earlier work, we provide — and incorporate into the optimization framework — rigorous error bounds that ensure reliability. In particular, in our optimization procedure we can (i) restrict attention to regions of the design space in which the model is provably accurate (or, alternatively, request adaptive improvement of the approximation as demanded by the optimization and design process), (ii) assess and control the suboptimality induced by the output approximation in the optimization results, and (iii), most importantly, rigorously ensure feasibility of the optimizers with respect to the exact mathematical description. The latter is particularly critical in real-time applications — in order to guarantee “safe” operation without recourse to (non-real-time) fiducial calculations.

There are two main components to our approach: the reduced-basis approximation (and error bound) which replaces the classical finite element approximation; and the appropriate incorporation of this approximation into the optimization procedure. The former are manifested in the latter in two important ways: in model constraints that, based on our error estimators, restrict attention to regions of the design space in which the approximation may be trusted — trusted to provide good predictions, and to provide meaningful gradients and Hessians; and in performance constraints that, through our error estimators, ensure that all proposed optimizers are in fact feasible with respect to the exact (more precisely, a very high-order “truth” finite element approximation of the exact) solution of the partial differential equation.

We do identify certain recommendations regarding the optimization procedure, and we provide the associated computational foundations. First, we recommend, given the smoothness of our problems (and approximations), that higher order information should be exploited; we thus develop efficient techniques for calculating the gradients and Hessians of our reduced-basis outputs and associated error estimators. Second, we recommend that the optimization procedure provide (intermediate) iterates that are strictly input-feasible, since violation of our “design-fidelity” constraints can lead to rapid divergence. And third, we recommend that, given the non-convex nature of our problems, the optimization procedure correctly avoid non-optimal stationary (saddle) points, and hence provide true (at least local) optimizers. An example that integrates all these components is described in [23].

References


Personnel Supported

Anthony T. Patera (principal investigator)  Professor, MIT Cambridge, MA
Martin A. Grepl  Graduate Student, MIT, Cambridge, MA
Ivan B. Oliveira  Postdoc, MIT, Cambridge, MA
Christophe Prud’homme  Postdoc, MIT, Cambridge, MA
Shu Heng (George) Pau  Graduate Student, MIT, Cambridge, MA
Karen Veroy  Postdoc, MIT, Cambridge, MA

Interactions

Other collaborators

Nguyen Ngoc Cuong (National University of Singapore and the Singapore-MIT Alliance) and Professor Liu Gui-Rong (National University of Singapore) have participated in the inverse-problem aspects of our studies. Many mathematical aspects of this work were developed in collaboration with Professor Yvon Maday (Paris-VI) and Maxime Barrault (Paris-VI).

Conferences


Publications


