RICAM Special Semester on Optimization

Workshop 3
Optimization and Inversion under Uncertainty

Book of Abstracts

(c) Johann Steininger

Linz, Austria
November 11-15, 2019

Invited Talks

Optimal Design of Inverse Problems Under Model Uncertainty,
with Application to Subsurface Flow

Alen Alexanderian
North Carolina State University

Abstract

We consider inverse problems that seek to infer an infinite-dimensional parameter from measurement data observed at a set of sensor locations and from the governing PDEs. We focus on the problem of optimal placement of sensors that result in minimized uncertainty in the inferred parameter field. This can be formulated as an optimal experimental design problem. We present a method for computing optimal sensor placements for Bayesian linear inverse problems governed by PDEs with model uncertainties. Specifically, given a statistical distribution for the model uncertainties we seek to find sensor placements that minimize the expected value of the posterior covariance trace; i.e., the expected value of the A-optimal criterion. The expected value is approximated using sample averaging leading to an objective function consisting of a finite sum of trace operators and a sparsity inducing penalty. Minimization of this objective requires many PDE solves in each step, making the problem extremely challenging. We will discuss strategies for making the problem computationally tractable. These include reduced order modeling and exploiting low-dimensionality of the measurements in the problems we target. We present numerical results for inference of the initial condition in a subsurface flow problem with inherent uncertainty in the velocity field.

This is joint work with Karina Koval (NYU) and Georg Stadler (NYU).

Risk-Averse Control of Fractional Diffusion and
a Gibbs Posterior Framework for Inverse Problems

Harbir Antil
George Mason University

Abstract

In the first part of this talk, we introduce and analyze a new class of optimal control problems constrained by elliptic equations with uncertain fractional exponents. We utilize risk measures to formulate the resulting optimization problem. We develop a functional analytic framework, study the existence of solution and rigorously derive the first-order optimality conditions. Additionally, we employ a sample-based approximation for the uncertain exponent and the finite element method to discretize in space.

In the second part of the talk, we adopt a general framework based on the Gibbs posterior to update belief distributions for inverse problems governed by partial differential equations (PDEs). The Gibbs posterior formulation is a generalization of standard Bayesian inference that only relies on a loss function connecting the unknown parameters to the data. It is particularly useful when the true data generating mechanism (or noise distribution) is unknown or difficult to specify. The Gibbs posterior coincides with Bayesian updating when a true likelihood function is known and the loss function corresponds to the negative log-likelihood, yet provides subjective inference in more general settings.

Expansions of Random Fields and Approximation of Uncertain PDEs

Markus Bachmayr
Johannes Gutenberg-Universität Mainz

Abstract

Many computational tasks in uncertainty quantification require random series representations of Gaussian random fields, which are typically derived by Karhunen-Loève expansions. However, alternative expansions are possible, and for a class of Gaussian random fields including those with Matérn covariances, we obtain expansions in terms of functions with wavelet-type multilevel structure. This hierarchical localization has implications on the computational costs of approximating PDEs with corresponding random coefficients. The construction of these wavelet-type representations is also closely connected to efficient methods for sampling random fields. This talk gives an overview of recent results in these directions.

Projected Stein Variational Newton:
A Fast and Scalable Bayesian Inference Method in High Dimensions

Peng Chen
Oden Institute of Computational Engineering and Science, University of Texas at Austin

Abstract

In this talk, we present a projected Stein variational Newton (pSVN) method for high-dimensional Bayesian inference. To address the curse of dimensionality, we exploit the intrinsic low-dimensional geometric structure of the posterior distribution in the high-dimensional parameter space via its Hessian (of the log posterior) operator and perform a parallel update of the parameter samples projected into a low-dimensional subspace by an SVN method. The subspace is adaptively constructed using the eigenvectors of the averaged Hessian at the current samples. We demonstrate fast convergence of the proposed method, complexity independent of the parameter and sample dimensions, and parallel scalability.

This is a joint work with Keyi Wu, Joshua Chen, Thomas O'Leary-Roseberry, and Omar Ghattas

Posterior Consistency for Bayesian Inverse Problems with Exponential Priors

Masoumeh Dashti
University of Sussex

Abstract

We consider the problem of recovering an unknown function from noisy and indirect observations. We adopt a Bayesian approach and consider a class of product exponential priors with tails between Gaussian and Laplace. We show some results on characterisation of the modes of the posterior measure, and under the assumption that the observations are generated from a fixed underlying value of the unknown, we discuss the consistency of the posterior modes and the rates of contraction of the posterior around this underlying value.

This is based on joint work with Sergios Agapiou (University of Cyprus) and Tapio Helin (LUT University).

Hierarchical Methods for Bayesian Inverse Problems

Matthew Dunlop
Courant Institute of Mathematical Sciences, New York University.

Abstract

Priors in Bayesian inverse problems may be endowed with additional flexibility via the use of hyperparameters; when the unknown is a function, they could for example allow for control over regularity, length-scale, or locations of jumps. A hierarchy of multiple layers of hyperparameters could also be used to allow for more complex structure. The posterior is then on both the unknown and these hyperparameters. In this talk I will overview different hierarchical structures appropriate for use in Bayesian inversion, wherein the unknown is typically high- or infinite-dimensional, and methods to robustly perform the inversion in practice. This includes methods of inferring the full posterior on the hyperparameters, and finding point estimates of them via MLE, MAP estimation or empirical Bayesian methods. We will also the consistency of such estimators: if a 'true' hyperparameter exists, can it be recovered in the limit of perfect data?

A Machine Learning Approach for Explicit Bayesian Inversion

Martin Eigel
WIAS Berlin

Abstract

Bayesian inverse problems can be solved explicitly, i.e. without sampling, yielding a functional approximation of the posteriori density, when a functional representation of the forward model is available. However, global chaos polynomials commonly used in UQ do not provide a favourable basis to accurately approximate highly concentrated densities. Instead, we propose a coordinate change which shifts the fast decay from the center of the density to a single dimension. This leads to a generalised polynomial chaos approximation which is h-refined just in a single variable. A variational formulation in the framework of statistical learning is employed to compute a low-rank density representation in a hierarchical tensor format. With the devised construction, the multi-element refinement leads to only a linear increase of the total number of basis functions.

The approach is illustrated with a Bayesian upscaling technique performed on a microstructured material, resulting in a stochastic description of a coarse effective material. From this, the constructed functional posterior density can directly be used for further numerical simulations in a stochastic Galerkin scheme.

This is joint work with Manuel Marschall (WIAS), Robert Gruhlke (WIAS) and Philipp Trunschke (TU Berlin).

Robust Uncertainty Propagation for Elliptic Diffusion Equations

Oliver Ernst
TU Chemnitz

Abstract

For the elliptic diffusion equation as the standard model problem for PDEs with random data, we investigate the sensitivity of several quantities with respect to perturbations in the input probability measure, among these the probability distribution of the solution, that of quantities of interest of the solution as well as risk functionals applied to such quantities of interest. We prove Lipschitz continuity with respect to total variation as well as Wasserstein distance. In contrast with previous sensitivity analyses, ours is based not on the perturbation of the uncertain inputs as random variables (fields), but on their underlying probability distributions, recalling that random variables with the same distribution may have positive distance in Lebesgue norms. We also show how to compare risk functionals of random variables evaluated with respect to different probability measures.

This is joint work with Alois Pichler (TU Chemnitz) and Björn Sprungk (U Göttingen).

Fast Methods for Bayesian Inverse Problems Governed by Random PDE Forward Models

Omar Ghattas
Oden Institute, The University of Texas at Austin, USA

Abstract

We consider Bayesian inverse problems governed by forward PDEs with random parameters. Specifically, given a prior probability of the inversion parameter m, a statistical model of observations d, and a forward model in the form of PDEs with a random parameter field k (representing model uncertainty), we wish to find the posterior probability of the inversion parameter. Besides the usual difficulties of Bayesian inversion, the random PDE forward model creates significant additional challenges. Finding the MAP point alone is a PDE-constrained optimization under uncertainty problem, and fully characterizing the posterior formally requires nested Monte Carlo sampling in random parameter and inversion parameter spaces. For complex forward problems, the inverse problem is thus intractable.

To address this challenge, we linearize the random-parameter-to-observable map, which leads to an explicit formula for the likelihood. New terms, involving the Fréchet derivative (J) of d with respect to k, appear in the likelihood, in particular in the noise covariance operator, where they represent model uncertainty. The explicit computation of J is prohibitive, requiring as many PDE solves as the lesser of the data and random parameter dimensions. Instead, a low rank approximation of J can be made efficiently via randomized SVD, typically requiring a small and (k) dimension-independent number of PDE solves.

Finding the MAP point (with respect to m) then requires solving a deterministic many-PDE-constrained optimization problem, with an objective involving the log posterior (enhanced with terms involving J), and with PDE constraints representing the forward problem, the generalized SVD problem for J, and incremental forward and adjoint problems representing actions of J and J^*. The gradient of the objective with respect to m can be found efficiently via an adjoint method, leading to an efficient MAP point solver. The Hessian of this functional evaluated at the MAP point approximates the posterior covariance.

Examples involving a subsurface flow problem and an advection–diffusion–reaction problem demonstrate both the need to account for model uncertainty (if only to first order) and the tractability of doing so.

This work is joint with Ilona Ambartsumyan (Oden Institute, The University of Texas at Austin).

Shape Optimization Under Uncertainty

Helmut Harbrecht
Departement of Mathematics and Computer Science, University of Basel, Switzerland

Abstract

Shape optimization is indispensable for designing and constructing industrial components. Many problems that arise in application, particularly in structural mechanics and in the optimal control of distributed parameter systems, can be formulated as the minimization of functionals which are defined over a class of admissible domains.

Shape optimization problems can be solved by means of gradient based minimization algorithms, which involve the shape functionals' derivative with respect to the domain under consideration. The computation of the shape gradient and the implementation of appropriate numerical optimization algorithms is meanwhile well understood, provided that the state equation's input data are given exactly. In practice, however, input data for numerical simulations in engineering are often not exactly known. One must thus address how to account for uncertain input data in the state equation.

Uncertainty in the state equation might arise from three different sources:

- Uncertainty might arise from geometric entities like a part of the boundary which has not to be optimized but is prescribed.

- The right-hand side of the state equation might be random.
- The material parameters, entering the partial differential operator, might be not exactly known.

We separately consider these sources of uncertainty and discuss their impact on the shape optimization problem. Especially, we show the well-posedness of the problem formulations and present numerical solution methods.

Probabilistic Constraints in Optimization with PDEs

René Henrion
Weierstrass Institute Berlin, Germany

Abstract

Probabilistic constraints offer a possibility to deal with uncertain parameters in the inequality constraints of some optimization problem. They provide decisions which are safe in a probabilistic sense given the distribution of those parameters. Originating from operations research they attract a growing recent interest in PDE constrained optimization in parallel with alternative approaches based on risk measures. The talk discusses some aspects of structural analysis (semicontinuity, convexity, existence of solutions) and algorithmic approaches (spheric-radial decomposition) for probabilistic programming in the context of infinite-dimensional decision variables. It then focuses on a concrete application to optimal Neumann boundary control of a vibrating string under uncertain initial conditions.

A Multilevel Stochastic Gradient Algorithm for PDE-Constrained Optimal Control Problems Under Uncertainty

Fabio Nobile
CSQI-MATH Ecole Polytechnique Fédérale de Lausanne, Switzerland

Abstract

We consider an optimal control problem for an elliptic PDE with random coefficients. The control function is a deterministic, distributed forcing term that minimizes an expected quadratic regularized loss functional. For its numerical treatment we propose and analyze a multilevel stochastic gradient (MLSG) algorithm which uses at each iteration a full, or randomized, multilevel Monte Carlo estimator of the expected gradient, build on a hierarchy of finite element approximations of the underlying PDE. The algorithm requires choosing proper rates at which the finite element discretization is refined and the Monte Carlo sample size increased over the iterations. We present complexity bounds for such algorithm. In particular, we show that if the refinement rates are properly chosen, in certain cases the asymptotic complexity of the full MLSG algorithm in computing the optimal control is the same as the complexity of computing the expected loss functional for one given control by a standard multilevel Monte Carlo estimator.

This is a joint work with Matthieu Martin (CRITEO, Grenoble), Panagiotis Tsilifis (EPFL), Sebastian Krumscheid (RWTH Aachen).

Risk Aversion in Dynamic Optimization

Alois Pichler
TU Chemnitz

Abstract

The dynamic programming principle is well investigated from mathematical perspective. The Hamiltonian drives the evolution equations, solutions are most often characterized by viscosity solutions. Its stochastic version typically involves the expected value.

This talk intends to extend the classical setting by incorporating risk in the objective, a risk averse assessment replaces the risk neutral expectation. Parametrizing risk aversion properly in time allows for a non-degenerate and risk averse mathematical setup. The corresponding evolution equation has a modified Hamiltonian, with a new term pointing in the direction of risk.

Explicit solutions are available for economic problems as Merton's fraction, which allows a nice interpretation of risk in addition.

Multistage Robust Convex Optimization Problems: A Sampling Based Approach

Georg Ch. Pflug
University of Vienna

Abstract

We consider multistage robust linear or convex optimization problems, which are - due to the typically infinite number of constraints - in general not accessible for numerical solution algorithms. We use random approximations of these problems, where just a finite number of constraints is randomly sampled and the problem is solved for this approximation. In multistage situations, this means that a random tree is generated and the problem is solved on this "scenario tree". Calafiore, Campi and Garatti have introduced the notion of "violation probability" for single stage situations indicating the probability that a newly independently generated constraint will leads to a different solution.

We give bounds for the violation probability at each stage of the decision process. The results are illustrated in an numerical example. It turns out that the empirical violation probability is much less than the bound given by the "worst case" argument. We may also determine necessary sample sizes to bring the violation probability under a given level. As a side result, we can prove the almost sure convergence of the random solution to the true non-random solution as the sample size (the bushiness of the tree) tends to infinity.

This is joint work with Fabrizio Dabbene (Torino) and Francesca Maggioni (Bergamo).

- F. Dabbene, F. Maggioni, G. Pflug. Multistage robust convex optimization problems: A sampling based approach. submitted.
- G. Calafiore. Random convex programs (2010). SIAM J. Optimization 20(6) 3427-3464
- M.C. Campi, S. Garatti. The exact feasibility of randomized solutions of uncertain convex programs (2008). SIAM J. Optimization 19 (3) 1211-1230.
- P. Vayanos, D. Kuhn, B. Rustem (2012). A constraint sampling approach for multi-stage robust optimization. Automatica 48 (3) 459-471.

Consistent Approximations in Optimization

Johannes O. Royset
Operations Research Department, Naval Postgraduate School

Abstract

Approximation is central to many optimization problems and the supporting theory provides insight as well as foundation for algorithms. In this talk, we lay out a broad framework for quantifying approximations by viewing finite- and infinite-dimensional constrained minimization problems as instances of extended real-valued lower semicontinuous functions defined on a general metric space. Since the Attouch-Wets distance between such functions quantifies epi-convergence, we are able to obtain estimates of optimal solutions and optimal values through bounds of that distance. In particular, we show that near-optimal and near-feasible solutions are effectively Lipschitz continuous with modulus one in this distance. Applications in optimal control and nonparametric statistics illustrate the theoretical results.

Stein Variational Newton and other Sampling-Based Inference Methods

Robert Scheichl
Heidelberg University

Abstract

High- or infinite-dimensional inverse problems in the context of complex physical systems arise in many science and engineering applications. Sampling-based inference methods provide an approach that allows in principle to solve this problem without suffering from the curse of dimensionality. However, bring these sampling methods in the feasible range for practical problems it is in general necessary to improve classical sampling approaches, such as random-walk Metropolis-Hastings MCMC. Stein variational gradient descent (SVGD) was recently proposed as a general purpose nonparametric variational inference algorithm: it minimizes the KullbackLeibler divergence between the target distribution and its approximation by implementing a form of functional gradient descent on a reproducing kernel Hilbert space [Liu & Wang, NIPS 2016]. In the main part of the talk, I will present a way how to accelerate and generalize the SVGD algorithm by including second-order information, thereby approximating a Newton-like iteration in function space. We also show how second-order information can lead to more effective choices of kernel. We observe significant computational gains over the original SVGD algorithm in multiple test cases. In the final part of the talk, I will present a few alternative approaches for accelerating sampling-based inference methods, such as low-rank tensor surrogates and multilevel ideas.

The talk presents joint work with G. Detommaso and S. Dolgov (Bath), T. Cui (Monash), C. Fox (Otago), Y. Marzouk and A. Spantini (MIT).

Algebraic Reasoning in Stochastic Optimization

Rüdiger Schultz
University of Duisburg-Essen

Abstract

Modeling, structural investigation, and algorithm design in traditional stochastic optimization, to large extent, rest on the one or other statement from mathematical analysis, with convexity in a prominent role.

The talk is devoted to parts of stochastic optimization where the theoretical underpinning comes from algebra. Ideal theory in rings of polynomials turns out fruitful when exploring feasibility, optimality and computation of optimal solutions to two- and multi-stage linear stochastic integer programs. When applied in this context, (Gröbner, Graver) standard bases undergo decomposition leading to a completely new class of algorithms.

Another relevant field in this respect is optimization with polynomials forming both the objectives and, via level sets, the constraints. This occurs, for instance, in steady-state models of potential-driven network flows. such as gas, power, and water. It results in properties of systems of polynomial equations that deserve further investigation.

Shape Optimization for Interface Identification in Nonlocal Models

Volker Schulz
Trier University, Germany

Abstract

Shape optimization methods have been proven useful for identifying interfaces in models governed by partial differential equations. Here we consider a class of shape optimization problems constrained by nonlocal equations which involve interface-dependent kernels. We derive a novel shape derivative associated to the nonlocal system model and solve the problem by a shape optimization algorithm, which is based on a Steklov-Poincaré-type shape metrics.

This is joint work with Christian Vollmann (Trier University).

Deep Neural Network Posterior Expression Rate Analysis
for Forward and Bayesian Inverse PDE UQ

Christoph Schwab
SAM, DMATH, ETH Zürich

Abstract

For well-posed operator equations subject to "distributed uncertain input data" from function spaces we analyze the expression rates by deep neural networks of several maps which arise in optimization and in uncertainty quantification (UQ). Adopting suitable (Riesz- or Schauder) bases representations of the uncertain input data converts to a contably-parametric, deterministic parametric optimization problem.

Examples comprise elliptic and parabolic fractional PDEs with uncertain coefficients and in uncertain domains, subject to uniform, Besov, and gaussian prior measures.

For forward PDE constrained UQ, we show (by bounding the DNN approximation error) that deep artificial ReLU networks express parametric solution manifolds without the curse of dimensionality.

For Bayesian inverse UQ, we show exponential expressivity of deep ReLU NNs for the data-to-QoI map, and for the Bayesian posterior density.

We also address the DNN generalization error.

New research directions will be outlined.

Joint work with Lukas Herrmann and Joost Opschoor (ETH) and Jakob Zech (MIT).

[1] Ch. Schwab and J. Zech: Deep learning in high dimension: Neural network expression rates for generalized polynomial chaos expansions in UQ, Analysis and Applications, Singapore **17/1** (2019), pp. 19-55.

[2] L. Herrmann and Ch. Schwab and J. Zech: Uncertainty Quantification for Spectral Fractional Diffusion: Sparsity Analysis of Parametric Solutions, SIAM/ASA J. Uncertainty Quantification **7/3** (2019), pp. 913-947.

[3] J. A. A. Opschoor and Ch. Schwab and J. Zech: Exponential ReLU DNN expression of holomorphic maps in high dimension, Report 2019-35, SAM, ETH Zurich `https://math.ethz.ch/sam/research/reports.html?id=839`

[4] L. Herrmann and M. Keller and Ch. Schwab: Quasi-Monte Carlo Bayesian estimation under Besov priors in elliptic inverse problems, Report 2019-41, SAM, ETH Zurich `https://math.ethz.ch/sam/research/reports.html?id=845`

On the Convergence of the Laplace Approximation and Noise-Level-Robustness of Laplace-Based Monte Carlo Methods for Bayesian Inverse Problems

Björn Sprungk
Institute for Mathematical Stochastics, Georg-August University of Göttingen

Abstract

The Bayesian approach to inverse problems provides a rigorous framework for the incorporation and quantification of uncertainties in measurements, parameters and models. We are interested in designing numerical methods which are robust w.r.t. the size of the observational noise, i.e., methods which behave well in case of concentrated posterior measures. The concentration of the posterior is a highly desirable situation in practice, since it relates to informative or large data. However, it can pose a computational challenge for numerical methods based on the prior or reference measure. We propose to employ the Laplace approximation of the posterior as the base measure for numerical integration in this context. The Laplace approximation is a Gaussian measure centered at the maximum a-posteriori estimate and with covariance matrix depending on the logposterior density. We discuss convergence results of the Laplace approximation in terms of the Hellinger distance and analyze the efficiency of Monte Carlo methods based on it. In particular, we show that Laplace-based importance sampling and Laplace-based quasi-Monte-Carlo methods are robust w.r.t. the concentration of the posterior for large classes of posterior distributions and integrands whereas prior-based importance sampling and plain quasi-Monte Carlo are not.

This is joint work with Claudia Schillings (University of Mannheim) and Philipp Wacker (FAU).

A New Primal-Dual Approach for Solving Risk-Averse PDE-Constrained Optimization Problems

Thomas M. Surowiec
Philipps-Universität Marburg

Abstract

Many stochastic optimization problems can be abstractly rewritten in the form

$$\min_{x \in \mathcal{X}_{\mathrm{ad}}} g(x) + \Phi(G(x)),$$

where $\mathcal{X}_{\mathrm{ad}}$ is a closed convex subset of a real Hilbert space $\mathcal{X}$, $g : \mathcal{X} \to \mathbb{R}$ is a continuously differentiable functional, $G : \mathcal{X} \to \mathcal{Y}$ is a nonlinear operator mapping $\mathcal{X}$ into another Hilbert space $\mathcal{Y}$, and $\Phi : \mathcal{Y} \to \mathbb{R}$ is a positively homogenous, subadditive, and monotone functional. In particular, risk-averse PDE-constrained optimization problems using coherent measures of risk as their risk models can often be rewritten in this way.

Starting with this abstract framework, we propose an algorithm that will ultimately be suitable for finite and infinite-dimensional optimization problems under uncertainty. The algorithm exploits the rich structure of convex risk measures. More specifically, we derive the steps of the algorithm from several observations on the application of the classical method of multipliers to the particular problem class. Using recent results on a variational smoothing technique for nonsmooth convex risk measures known as epi-regularization, we provide convergence results for both convex and nonconvex problem settings. The performance of the algorithm is demonstrated via several examples from the literature and compared to a well-known solver for non-smooth optimization problems.

This is joint work with Drew P. Kouri (Sandia National Laboratories).

An Approximation Scheme for Distributionally Robust PDE-Constrained Optimization

Michael Ulbrich
Technical University of Munich

Abstract

We develop a sampling-free approximation scheme for distributionally robust optimization (DRO) with PDEs. The DRO problem can be written in a bilevel form that involves maximal (i.e., worst case) value functions of expectations of nonlinear functions that depend on the optimization variables and random parameters. The maximum values are taken over an ambiguity set of probability measures which is defined by moment constraints. To achieve a good compromise between tractability and accuracy we approximate nonlinear dependencies of the cost / constraint functions on the random parameters by quadratic Taylor expansions. This results in an approximate DRO problem which on the lower level then involves value functions of parametric trust-region problems and of parametric semidefinite programs. Using trust-region duality, a barrier approach, and other techniques we construct gradient consistent smoothing functions for these value functions and show global convergence of a corresponding homotopy method. We discuss the application of our approach to PDE constrained optimization under uncertainty and present numerical results.

This is joint work with Johannes Milz.

Robust Optimization Approaches for PDE-Constrained Optimization Under Uncertainty

Stefan Ulbrich
Department of Mathematics, TU Darmstadt

Abstract

We consider robust nonlinear PDE-constrained optimization problems with ellipsoidal uncertainty sets yielding a min-max formulation. Using second-order Taylor approximations (we also discuss briefly first order approximations) of the objective and constraint functions w.r.t. uncertain parameters, the inner maximization problems are reduced to trust-region subproblems. We consider two solution methods for the approximated robust counterpart: a nonsmooth formulation using the maximum-value functions of the inner problems, and an MPCC formulation. We propose an update procedure for the expansion point of the Taylor model w.r.t. uncertain parameters during the optmization. Moreover, we discuss the incorporation of reduced order models with error estimation within robust optimization methods as well as the usage of iterative trust region-solvers to increase the computational efficiency. We apply the method to shape optimization for electrical engines and mechanical structures.

This is joint work with Philip Kolvenbach, Oliver Lass, Herbert De Gersem, Sebastian Schöps and in parts with Michael Hinze and Allessandro Alla.

On Multilevel Best Linear Unbiased Estimators

Elisabeth Ullmann
Department of Mathematics, Technical University of Munich

Abstract

We discuss novel multilevel best linear unbiased estimators (BLUEs) introduced in [4]. The goal is the estimation of the expectation of a scalar-valued quantity of interest associated with a family of multifidelity models. The key idea of the multilevel BLUE is to reformulate the estimation as a linear regression problem. By construction, BLUEs are variance minimal within the class of linear unbiased estimators. We compare our proposed estimator to other multilevel estimators such as multilevel Monte Carlo [1], multifidelity Monte Carlo [3], and approximate control variates [2]. In addition, we show that our estimator approaches a sharp lower bound that holds for any linear unbiased multilevel estimator in the infinite low-fidelity data limit. Finally, we specialize the results in [4] to PDE-based models which are parameterized by a discretization quantity, e.g. the finite element mesh size. We prove that in this case, the complexity of the BLUE is not worse than the complexity of multilevel Monte Carlo. In practise, we observe a complexity reduction for selected random elliptic PDE problems.

References

[1] M. B. Giles, *Multilevel Monte Carlo methods*, Acta Numerica, 24, pp. 259–328, 2015.

[2] A. A. Gorodetsky, G. Geraci, M. Eldred, J. D. Jakeman, *A Generalized Framework for Approximate Control Variates*, Preprint arXiv:1811.04988, 2018.

[3] B. Peherstorfer, K. Willcox, M. Gunzburger, *Optimal Model Management for Multifidelity Monte Carlo Estimation*, SIAM J. Sci. Comput., 38, pp. A3163–A3194, 2016.

[4] D. Schaden, E. Ullmann, *On multilevel best linear unbiased estimators*, Preprint No. IGDK-2019-11, May 2019. Submitted.

[5] D. Schaden, E. Ullmann, *Asymptotic analysis of multilevel best linear unbiased estimators*. Submitted.

Multilevel Monte Carlo Methods for the Robust Optimization of Systems Described by Partial Differential Equations

Stefan Vandewalle

Department of Computer Science, KU Leuven - University of Leuven

Abstract

We consider PDE-constrained optimization problems, where the partial differential equation has uncertain coefficients modelled by means of random variables or random fields. The goal of the optimization is to determine an optimum that is satisfactory in a broad parameter range, and as insensitive as possible to parameter uncertainties. First, an overview is given of different deterministic goal functions which achieve the above aim with a varying degree of robustness. Next, a multilevel Monte Carlo method, see [1], is presented which allows the efficient calculation of the gradient and the Hessian arising in the optimization method. The convergence and computational complexity for different gradient and Hessian based optimization methods is then illustrated for a model elliptic diffusion problem with lognormal diffusion coefficient [2]. We also explain how the optimization algorithm can benefit from taking optimization steps at different levels of the multilevel hierarchy, in a classical MG/OPT framework [3]. We demonstrate the efficiency of the algorithm, in particular for a large number of optimization variables and a large number of uncertainties.

1. M. B. Giles, Multilevel Monte Carlo Methods. Acta Numerica, 24, pp. 259–328, 2015.
2. A. Van Barel, S. Vandewalle. Robust Optimization of PDEs with Random Coefficients Using a Multilevel Monte Carlo Method. SIAM/ASA Journal on Uncertainty Quantification 7 (1), pp. 174–202, 2019.
3. S.G. Nash. A Multigrid Approach to Discretized Optimization Problems. Optimization Methods and Software, 14 (1-2), pp. 99–116, 2000.

Posters

A Primal-Dual Algorithm for PDE-Constrained Optimization Involving the Conditional Value-at-Risk

Sebastian Angerhausen
University of Duisburg-Essen

Abstract

We consider a stochastic, risk-averse optimization problem in an infinite-dimensional Hilbert space. The objective function includes the Conditional Value-at-Risk (CVaR) as a risk measure, as well as the solution operator of a PDE-constraint. We derive the optimality conditions and formulate the proximal point mapping of the Fenchel conjugate of CVaR as the projection onto the bounded probability simplex. This projection is easy to compute once the problem has been discretized. The presented algorithm is adapted from the first-order primal-dual algorithm introduced by A. Chambolle and T. Pock in 2010, together with the extension by T. Valkonen in 2014, which allows the use of nonlinear operators. Numerical experiments are carried out after discretization and sampling from the underlying probability distribution. The step sizes are chosen individually for each sample and adapted during the runtime of the algorithm.

Multi-Level Monte Carlo Estimators for Gradient-Based Optimisation in Engineering Design

Quentin Ayoul-Guilmard
CSQI, École Polytechnique Fdrale de Lausanne, Switzerland

Abstract

Robustness to meteorological conditions is an important factor in civil engineering and motivates risk-averse approaches to shape design. Such optimal shapes can generally be defined with a problem of optimisation under uncertainties (OUU) featuring a risk measure dependent on the solution of a random boundary-value problem. This OUU problem can be solved in a number of ways, amongst which are gradient-based methods requiring computations of the sensitivities of the objective function with respect to the design parameters. Therefore, an accurate and efficient estimation of these statistics is paramount.

We present methods to estimate such statistics for risk measures expressible as parametric expectations (e.g. mean, quantile, conditional-value-at-risk) with multi-level Monte Carlo estimators. We show efficient algorithms adaptively calibrated from error indicators. Finally, we investigate and discuss their use in gradient-based optimisation techniques.

Solving a Bernoulli Type Free Boundary Problem with Random Diffusion

Rahel Brügger
University of Basel

Abstract

We consider Bernoulli's exterior free boundary problem with random diffusion. For its numerical solution we reformulate it as a shape optimization problem and seek to minimize the ensemble average of the random shape functional. The domain under consideration is represented by a level set function which is evolved by the objective's shape gradient. The state is computed by the finite element method, where the underlying triangulation is constructed by means of a marching cubes algorithm. The random shape functional is approximated by the quasi-Monte Carlo method. Numerical experiments validate the feasibility of the approach. This is joint work with Roberto Croce and Helmut Harbrecht.

Projected Stochastic Gradient Method
for PDE Constrained Optimization Under Uncertainty

Caroline Geiersbach
University of Vienna

Abstract

This research is concerned with efficient stochastic algorithms for solving problems in PDE constrained optimization under uncertainty. The approach is based on the stochastic gradient algorithm originating from a pioneering paper by Robbins and Monro (1951). To adapt the projected stochastic gradient method to the problem class, we observe its convergence in Hilbert spaces. In particular, we produce convergence results for smooth convex objectives with closed and convex constraints. We show that the method produces a sequence that converges in probability to the solution set, even in the presence of numerical error.

The method is adapted to a model problem, the optimal control of a stationary heat source with box constraints, where uncertainty arises from the material coefficient, which is modeled as a random field. Additionally, numerical error arising from finite element discretization of the PDE constraint is analyzed, taking into account the regularity of the random field. Efficiency estimates for the expected error in the objective function and iterates are derived, which incorporate numerical error and cover both the strongly convex case and the general convex case with iterate averaging. Based on these estimates, a mesh refinement strategy is introduced to balance numerical error with decreasing step sizes. Numerical experiments confirm the expected convergence rates.

A Quasi-Monte Carlo Method for PDE-Constrained Optimization Under Uncertainty

Philipp Guth
University of Mannheim

Abstract

We apply a quasi-Monte Carlo (QMC) method to an optimal control problem constrained by an elliptic partial differential equation (PDE) equipped with an uncertain diffusion coefficient. In particular, the optimization problem is to minimize the expected value of a tracking type cost functional with an additional penalty on the control. The uncertain coefficient in the PDE is parametrized via a Karhunen–Loève (KLE) expansion and the expected value is considered as an infinite-dimensional integral in the corresponding parameter space. We discretize the optimization problem by truncating the KLE after s terms, approximating the expected value by an n-point QMC rule and approximating the solution of the PDE using finite elements (FE). We show that the discretization error of the solution to the optimization problem is bounded by the discretization error of the adjoint state. For the convergence analysis the latter is decomposed into truncation error, QMC error and FE error.

This is joint work with Vesa Kaarnioja (UNSW Sydney, Australia), Frances Kuo (UNSW Sydney, Australia), Claudia Schillings (University of Mannheim, Germany), Ian Sloan (UNSW Sydney, Australia).

On the Well-Posedness of Bayesian Inverse Problems

Jonas Latz
Technische Universität München

Abstract

The subject of this presentation is a new concept of well-posedness of Bayesian inverse problems. The conventional concept of (Lipschitz, Hellinger) well-posedness in [Stuart 2010, Acta Numerica 19, pp. 451-559] is difficult to verify in practice and may be inappropriate in some contexts. Our concept simply replaces the Lipschitz continuity of the posterior measure

in the Hellinger distance by continuity in an appropriate distance between probability measures. Aside from the Hellinger distance, we investigate well-posedness with respect to weak convergence, the total variation distance, the Wasserstein distance, and also the Kullback–Leibler divergence. We demonstrate that the weakening to continuity is tolerable and that the generalisation to other distances is important. We give well-posedness results with respect to some of the aforementioned distances for large classes of Bayesian inverse problems. Here, little or no information about the underlying model is necessary.

Infinite Swapping in Bayesian Inverse Problems

Juan Pablo Madrigal Cianci
École Polytechnique Fédérale de Lausanne

Abstract

Based on ideas originating in the molecular dynamics community, we present the infinite swapping algorithm on a Markov chain Monte Carlo (MCMC) context, which can be understood as a generalization of the by now standard parallel tempering algorithm. In such approach, instead of proposing to swap states between two of the K chains and accepting such swap using a Metropolis-Hastings rule, the infinite swapping approach proposes a a swap of dynamics (i.e, tempering parameter and Markov kernel) between all $K!$ possible combinations of such parameters at every step of the algorithm. This can be interpreted, from a continuous time perspective, as the process of swapping states infinitely often in the parallel tempering algorithm.

We introduce such sampling strategy in a Bayesian inverse problem context, particularly aimed at sampling from so-called non-identifiable posterior distributions. We implement such algorithm to sample from various posterior distributions, and compare it with more traditional MCMC techniques, such as parallel tempering and random walk Metropolis.

Newton-Based Methods for the Numerical Solution of Risk-Averse PDE-Constrained Optimization Problems

Mae Markowski
Rice University

Abstract

This poster introduces a modification of Newton-Conjugate Gradient (CG) type methods applied to smoothed Conditional-Value-at-Risk (CVaR) optimization problems. Furthermore, it presents efficient gradient and Hessian-times-vector computations for smoothed CVaR optimization governed by partial differential equation (PDE) constraints that use the CVaR structure to reduce the number of adjoint PDE and linearized PDE solves needed during optimization. Because the CVaR objective is non-smooth, smoothing is frequently applied to allow the application of Newton-type methods. However, rank deficiency in the Hessian of the smoothed CVaR objective arises because the second derivative of the smoothing function is zero outside a small interval. This rank deficiency leads to unbounded quadratic models, which yield large Newton-CG steps and subsequently produce small step-sizes or trust-region radii. The proposed modification introduces a computationally inexpensive sub-step, which utilizes the structure of the smoothed CVaR objective to generate a consistent quadratic model. This leads to substantially better Newton-CG steps and significantly reduces the number of Newton-CG iterations, which is demonstrated on a model problem arising from optimal control.

Intelligent Multifidelity Inference in Probabilistic Programming

Ludger Paehler
Technical University of Munich

Abstract

We propose a synergy of the multifidelity paradigm with recent advances in probabilistic programming and hierarchical reinforcement learning based sampling management. Building on Gen, a new probabilistic programming system with programmable inference, control-variate multifidelity Monte-Carlo is transferred into the probabilistic programming syntax. Rooted in Bayesian inference, the domain-specific language syntax of the probabilistic programming language affords us the benefits of syntactical constructs such as *choicemaps* and *regenerate* and a natural reasoning about uncertainty. Applying a novel implementation of the hierarchical proximal policy optimization we extend the exploitation of the hierarchical structure of our simulations from the inference framework to the training of our stochastic sampling policy. Bootstrapping the sampling policy from low-fidelity inference routines, transfer learning is applied to continuously adopt the policy to the optimal high-fidelity inference routine by exploiting the model hierarchy. Highly competitive performance is demonstrated on a suite of model problems.

This is joint work with Nikolaus Adams (TUM).

Pre-Asymptotic Recovery of Reduced Models from Data

Wayne Isaac Tan Uy
Courant Institute of Mathematical Sciences

Abstract

Model reduction is a well-established framework for reducing the simulation cost of large-scale dynamical systems while simultaneously offering rigorous error estimates for certain problem classes. However, traditional model reduction is intrusive in the sense that the full-model operators of the system are required. In contrast, we pursue a non-intrusive model reduction approach that infers the operators of the reduced model from data of the full model such as trajectories, inputs, residuals, and initial conditions. The inferred reduced model is identical to traditional model reduction if a reprojection scheme is performed on data of the full model to preserve Markovian dynamics in the reduced subspace. The data-driven reduced model therefore inherits well-studied properties such as error control. Numerical results show that the proposed approach is able to recover the reduced model arising from the traditional framework up to numerical errors using only finite amounts of data. This is joint work with Benjamin Peherstorfer.