ÖAW AUSTRIAN ACADEMY OF SCIENCES

RICAM
JOHANN · RADON · INSTITUTE
FOR COMPUTATIONAL AND APPLIED MATHEMATICS

RICAM Special Semester on Multivariate Algorithms and Their Foundations in Number Theory

Workshop 2
Multivariate Algorithms and Information-Based Complexity

Book of Abstracts

Linz, Austria
November 5–9, 2018

Off-axis anisotropy in multivariate functions

Paul Constantine
University of Colorado, Boulder

Abstract

Modern computational science often involves approximation or integration of multivariate functions arising from physics-based applications, where a function maps physical input parameters to output quantities of interest. We often observe that these functions have strong anisotropic dependence on the inputs—i.e., the outputs vary primarily along a handful of directions in the input space, and these directions rarely align with the input axes. I will show several examples of such anisotropic structures and discuss techniques for discovering and exploiting the anisotropy for approximation and integration.

The benefits of kernels on closed manifolds

Martin Ehler
University of Vienna

Abstract

This talk is intended to provide an overview and some insights into approximation problems on closed manifolds. We shall discuss the approximation of integrals, functions, and the approximation of continuous measures by discrete ones. These three approximation tasks refer to (pseudo-) Monte-Carlo theory, to classical function approximation, and to discrepancy theory. We shall discuss them within a joint framework provided by the theory of reproducing kernels. Numerical experiments for the sphere, the special orthogonal group, and the Grassmannian manifold support the theoretical results.

Lower error bounds for strong approximation of scalar SDEs with non-Lipschitzian coefficients

Mario Hefter
TU Kaiserslautern

Abstract

We study pathwise approximation of scalar stochastic differential equations at a single time point or globally in time by means of methods that are based on finitely many observations of the driving Brownian motion. We prove lower error bounds in terms of the average number of evaluations of the driving Brownian motion that hold for every such method under rather mild assumptions on the coefficients of the equation. The underlying simple idea of our analysis is as follows: the lower error bounds known for equations with coefficients that have sufficient regularity globally in space should still apply in the case of coefficients that have this regularity in space only locally, in a small neighborhood of the initial value. Our results apply to a huge variety of equations with coefficients that are not globally Lipschitz continuous in space including Cox-Ingersoll-Ross processes, equations with superlinearly growing coefficients, and equations with discontinuous coefficients. In many of these cases the resulting lower error bounds even turn out to be sharp.

This is joint work with André Herzwurm and Thomas Müller-Gronbach.

Complexity of stochastic integration in Sobolev spaces

Stefan Heinrich
TU Kaiserslautern

Abstract

We study stochastic integration in classes of functions in the framework of information-based complexity theory. We survey recent and present some new general results on lower bounds in the deterministic and randomized setting. These results are applied to stochastic integration problems in Sobolev classes. We also discuss algorithms and matching upper bounds, this way determining the complexity of these problems. This is related to recent work by Eisenmann and Kruse [1], who proposed randomized algorithms and obtained error estimates for stochastic integration of functions from fractional Sobolev spaces.

[1] Monika Eisenmann, Raphael Kruse, Two quadrature rules for stochastic Ito integrals with fractional Sobolev regularity, `arXiv:1801.08531`.

Local RBF-based penalized least-squares approximation of noisy scattered data on the sphere

Kerstin Hesse
University of Paderborn

Abstract

In geophysical applications, measured scattered data usually contains noise, and any approximation method should take this into account. Moreover, the data may be local rather than global, for example, if the data is collected by satellites, there is usually a polar gap (i.e. no data in a neighborhood of the poles).

In this talk we discuss the properties of a radial basis function (RBF) based penalized least-squares approximation of a local function on the unit sphere $\mathbb{S}^d$ of $\mathbb{R}^{d+1}$ from local noisy data. More precisely, the approximated function f is defined on a subset Ω of $\mathbb{S}^d$ with Lipschitz-continuous boundary, lies in the Sobolev space $W_2^s(\Omega)$ with $s > d/2$, and is given in the form of noisy data on a finite point set $X_N \subseteq \overline{\Omega}$. The approximation is the minimizer of a certain quadratic functional which depends on a smoothing parameter $\lambda > 0$ that balances between fitting the data and getting a smooth solution. This minimizer is a (globally defined) RBF approximation with the centers of the RBFs at the points of X_N. This RBF approximation is computed by solving a linear system with a positive definite matrix. The minimized functional can be interpreted as an instance of Tikhonov regularization.

A crucial question is how the smoothing parameter λ should be chosen depending on the noise level, and in this talk we consider one a posteriori strategy for choosing λ, namely Morozov's discrepancy principle. For λ chosen with Morozov's discrepancy principle, we present order-optimal L_2 error estimates in terms of powers of the mesh norm of X_N and the noise level. The proof of these L_2 error estimates involves a local L_2 sampling inequality as well as a suitable Sobolev extension theorem. A numerical example illustrates the theoretical work.

The presented results are from recent and ongoing joint work with Ian H. Sloan and Robert S. Womersley.

Adaptive approximation to multivariate linear problems for inputs lying in a cone

Fred J. Hickernell
Illinois Institute of Technology

Abstract

Function recovery, differentiation, and integral equations are examples of multivariate problems which require approximate numerical solutions. One would like to identify a good algorithm, analyze the computational cost to achieve the desired error tolerance, understand how this cost depends on the number of variables, and determine whether the proposed algorithm is nearly optimal relative to the best possible algorithm. This talk focuses on the situation where the inputs can be represented as series, and where the algorithm is allowed go sample individual series coefficients. Rather than focusing on inputs lying inside a ball of fixed radius, we focus on inputs lying inside a cone of nice inputs. This allows us to construct an adaptive algorithm, that automatically determines the number of series coefficients required to achieve the desired error tolerance. The computational cost of the algorithm is characterized in terms of the space of inputs, space of outputs, solution operator, and definition of the cone. The information-based complexity of problem and its tractability are also characterized. These depend on the relative importance of the variables.

High accuracy algorithms for interpolating and integrating multivariate functions defined by sparse samples in high dimensions

James (Mac) Hyman
Tulane University

Abstract

We will describe and analyze accurate and efficient numerical algorithms to interpolate and approximate the integral of multivariate functions. The algorithms can be applied when we are given the function values at an arbitrary positioned, and usually small, existing sparse set of function values (samples), and additional samples are impossible, or difficult (e.g. expensive) to obtain. The methods are based on local, and global, tensor-product sparse quasi-interpolation methods that are exact for a class of sparse multivariate orthogonal polynomials.

Although the algorithms can be applied for general distributions of sample points, they are especially effective in improving the approximation of integrals defined on low-discrepancy (approximately uniformly distributed) sample. One advantage of this approach is that it continues to be effective on even high-discrepancy sampling distributions. This allows the algorithms to be used with adaptive sampling algorithms that concentrate sample points in regions with locally high gradients.

We will compare the convergence rate of the method on smooth and discontinuous functions defined on both low discrepancy quasi-Monte Carlo distributions of samples, as well as for sample distributions that are far from uniformly distributed. We observe that although the approach does not change the error convergence rate, the constant multiplicative factor in the error is often reduced by a factor of over 100 or more.

On strong approximation of SDEs wth discontinuous drift

Thomas Müller-Gronbach
University of Passau

Abstract

We present recent results on pathwise approximation of SDEs with a discontinuous drift coefficent.

This is joint work with Larisa Yaroslavtseva (ETH Zürich and University of Passau).

How good is random information?

Erich Novak
University of Jena

Abstract

We study problems like approximation and integration and compare the quality (radius) of optimal information with the radius of random information.

For some problems random information is almost optimal and for some other problems random information is very bad compared with optimal information.

This is joint work with Aicke Hinrichs and David Krieg.

What is an optimal Bayesian method?

Chris Oates
University of Newcastle

Abstract

It is well-understood that Bayesian decision theory and average case analysis are essentially identical [1]. However, for the purposes of performing *uncertainty quantification* it can be argued that the Bayesian decision theoretic framework is neither appropriate nor sufficient [2]. To this end, we consider an alternative criterion for the design of a Bayesian method and compare its implied optimal information to the information used in an average case optimal method. In general the two are not identical and this suggests a new regime in which experimental design and information-based complexity can be studied.

[1] J.B. Kadane and G.W. Wasilkowski. Average Case ϵ-Complexity in Computer Science: A Bayesian View, in Bayesian Statistics, pages 361374. Elsevier, North-Holland, 1985.

[2] J. Cockayne, C. J. Oates, T. Sullivan and M. Girolami. Bayesian Probabilistic Numerical Methods. `arXiv:1702.03673`.

A game theoretic approach to numerical approximation, algorithm design and learning

Houman Owhadi
California Institute of Technology

Abstract

This talk will cover interplays between Game Theory, Numerical Approximation, Gaussian Process Regression and Learning. We will illustrate the interface between statistical inference and numerical analysis through problems related to numerical homogenization, operator adapted wavelets, fast solvers, computation with dense kernel matrices. We will also show how this perspective can be applied in machine learning to the design of bottomless networks (Kernel Flows) that are amenable to some degree of analysis (these networks simulate a data driven stochastic flow in the input space and learn kernels capable of generalization from one interpolation point per class).

This talk will cover joint work with F. Schäfer, C. Scovel, T. Sullivan, G.R. Yoo and L. Zhang.

On the regularized functional regression

Sergei Pereverzyev
Johann Radon Institute for Computational and Applied Mathematics,
Austrian Academy of Sciences

Abstract

Functional Regression (FR) involves data consisting of a sample of functions taken from some population. Most work in FR is based on a variant of the functional linear model first introduced by Ramsay and Dalzell in 1991. A more general form of polynomial functional regression has been introduced only quite recently by Yao and Mller (2010), with quadratic functional regression as the most prominent case. A crucial issue in constructing FR models is the need to combine information both accross and within observed functions, which Ramsay and Silverman (1997) called replication and regularization, respectively. In this talk we are going to present a general approach for the analysis of regularized polynomial functional regression of arbitrary order and indicate the possibility for using here a technique that has been recently developed in the context of supervised learning. Finally we are briefly discuss the application of FR in stenosis detection.

Joint research with S. Pereverzyev Jr. (Uni. Med. Innsbruck), A. Pilipenko (IMATH, Kiev) and V.Yu. Semenov (DELTA SPE, Kiev), supported by the consortium of Horizon-2020 project AMMODIT and the Austrian National Science Foundation (FWF).

Absolute Value Information from the IBC perspective

Leszek Plaskota
University of Warsaw

Abstract

Two classes of information have been mainly considered in IBC for approximate solutions of continuous problems: $\Lambda^{\rm all}$ that consists of all linear functionals, and $\Lambda^{\rm std}$ that consists of only function evaluations. A different class of information has been studied in the context of phase retrieval, where it is assumed that only absolute values of linear functionals from $\Lambda \subseteq \Lambda^{\rm all}$ are available. We denote it $|\Lambda|$ and call *the absolute value information.* For $|\Lambda|$ we need to modify the algorithm error to compensate the missing phase in information values.

The purpose of this study is to establish the powers of $|\Lambda^{\rm all}|$ and $|\Lambda^{\rm std}|$ in comparison to $\Lambda^{\rm all}$ and $\Lambda^{\rm std}$ for various IBC problems in the worst case setting. Our main result is that $|\Lambda^{\rm all}|$ is roughly of the same power as $\Lambda^{\rm all}$ for linear IBC problems. In fact, for the complex case this holds for all Λ with the property that $L_1, L_2 \in \Lambda$ implies that $L_1 + L_2$ and $L_1 + {\rm i}\, L_2$, with ${\rm i} = \sqrt{-1}$, are also in Λ. In general, this property is *not* satisfied by $\Lambda^{\rm std}$. We prove that $|\Lambda^{\rm std}|$ is usually too weak to solve linear problems.

This is a joint work with Paweł Siedlecki and Henryk Woźniakowski.

Path-dependent vs path-independent adaptive step-size control for jump-diffusion SDEs

Paweł Przybyłowicz
AGH University of Science and Technology in Kraków

Abstract

In the talk we discuss strong global approximation of solutions of stochastic differential equations (SDEs) of the following form

$$\begin{cases} dX(t) = a(t, X(t))dt + b(t, X(t))dW(t) + c(t, X(t-))dN(t), \quad t \in [0, T], \\ X(0) = x_0, \end{cases} \tag{1}$$

where $x_0 \in \mathbf{R}$, $a, b, c : [0,T] \times \mathbf{R} \to \mathbf{R}$ satisfy certain regularity conditions, $W = \{W(t)\}_{t\in[0,T]}$ is a one-dimensional Wiener process and $N = \{N(t)\}_{t\in[0,T]}$ is a homogeneous Poisson process with intensity $\lambda > 0$.

We consider three classes of methods χ^{eq}, χ^{noneq} and $\chi^{\mathrm{noneq}*}$. Algorithms from the class χ^{eq} use only equidistant sampling for $[W,N]$, methods from the second class χ^{noneq} rely on the nonequidistant but path-independent sampling for $[W,N]$, while algorithms that belong to $\chi^{\mathrm{noneq}*}$ might use adaptive (path-dependent) sampling with respect to trajectories of the driving processes $[W,N]$. In the latter case the sampling consists of sequence of stopping times. In each class we define suitable schemes, that are based on specific interpolation of the classical Milstein steps. Moreover, we establish their exact asymptotic errors. It turns out that the methods based on adaptive sampling with respect to the trajectories of $[W,N]$ are more efficient than those based on the equidistant/nonequidistant path-independent discretization. We also discuss optimality of the presented methods and report results of numerical experiments.

Random bits vs. random numbers for an infinite-dimensional quadrature problem

Klaus Ritter
TU Kaiserslautern

Abstract

Infinite-dimensional quadrature problems naturally arise in the context of stochastic differential equations (SDEs): If X is the solution of the SDE and φ is a real-valued functional on the path space, then the expectation of $\varphi(X)$ is an infinite-dimensional integral (with respect to the, typically, unknown distribution of X).

We study the approximation of such expectations by means of randomized (Monte Carlo) algorithms that may use either random numbers or merely random bits. To compare the power of random numbers and random bits we present upper and lower bounds on the complexity for the corresponding classes of randomized algorithms. The upper bounds are achieved by suitable multilevel methods.

This talk is based on joint work with M. Giles (Oxford), M. Hefter, and L. Mayer (both from Kaiserslautern).

Efficient kernel-based learning by localization

Ingo Steinwart
University of Stuttgart

Abstract

Despite the recent successes of (deep) neural networks kernel-based learning (KBL) methods remain one of the most successful learning methods for unstructured small to medium sized classification and regression problems. However, when it comes to large scale applications, their computational requirements, which grow super-linearly in the number of training samples, renders their application infeasible. To address this issue, several approaches that e.g. train KBL on many small chunks of the given large data set separately have been proposed in the literature. In this talk, we consider such a decomposition strategy, called localized KBL, which is based upon a spatial partition of the input space. For this localized KBL, we derive a general oracle inequality describing its learning performance. Then we apply this oracle inequality to least squares regression using Gaussian kernels and deduce local learning rates that are essentially minimax optimal under some standard smoothness assumptions on the regression function. We further introduce a data-dependent parameter selection method for our localized KBL approach and show that this method achieves the same learning rates as before. Finally, we present some larger scale experiments for our localized KBL showing that it achieves essentially the same test performance as a global KBL for a fraction of the computational requirements. In addition, it turns out that the computational requirements for the local KBLs are similar to those of a vanilla random chunk approach, while the achieved test errors are significantly better.

Euler-type schemes for SDEs with discontinuous drift

Michaela Szölgyenyi
Alpen-Adria Universität Klagenfurt

Abstract

Stochastic differential equations with irregular (non-globally Lipschitz) coefficients are a very active topic of research. We study equations, where we relax the global Lipschitz condition on the drift coefficient to allow for discontinuities on a set of positive reach.

We study strong convergence of an Euler-type scheme, which uses adaptive step-sizing for a better resolution close to the discontinuity.

We obtain a numerical method which has – up to logarithmic terms – strong convergence order 1/2 with respect to the average computational cost.

Joint work with A. Neuenkirch (University of Mannheim) and Lukasz Szpruch (University of Edinburgh).

∞-variate integration

G. W. Wasilkowski
University of Kentucky

Abstract

We present recent results on efficient approximation of integrals with infinitely many variables. We provide new concepts of *worst case truncation* and *superposition dimensions*. We show that for weighted tensor product spaces, these dimension are small for modest error demands.

We also present *Multivariate Decomposition Method (MDM)*. It is based on approximating a number of integrals, each with the number of variables bounded by the superposition dimension. Therefore approximating such integrals by *Sparse Grid* or *Quasi-Monte Carlo Methods* leads to very efficient algorithms.

These results are presented for rather general spaces and problems. Next, we apply them for the problem of approximating weighted integrals over unbounded domains for weighted spaces of functions with mixed partial derivatives of order one bounded in a weighted L_p norm. We propose a special transformation that allows well-known Sparse Grid or Quasi-Monte Carlo methods to be used with the optimal rate of convergence.

The results mentioned above are for *anchored* spaces. Finally, we present results on the equivalence of anchored and *ANOVA* spaces, which imply that efficient algorithms for anchored spaces are also efficient for ANOVA spaces.

The presentation is based on papers co-authored with A. Gilbert, M. Gnewuch, M. Hefter, A. Hinrichs, P. Kritzer, F. Y. Kuo, D. Nuyens, F. Pillichshammer, L. Plaskota, K. Ritter, I. H. Sloan, and H. Woźniakowski.

Tractability condtions for linear multivariate problems

Henryk Woźniakowski
Columbia University and University of Warsaw

Abstract

We present necessary and sufficient conditions on various notions of tractability for the exponential case for linear multivariate problems. They are presented in terms of eigenvalues of certain compact self-adjoint operator.

Joint work work with Peter Kritzer.

Algebraic methods in sparse grid interpolation

Henry P. Wynn
London School of Economics

Abstract

Sparse grids are used for both polynomial quadrature and interpolation, particularly combined with stochastic finite element methods under the modern heading of Uncertainty Quantification. The grids are typically permuted versions of a base grid whose form turns out to have a strong connection to the theory of monomial ideals. This allows the complex combinatorial constants in interpolation formulae, and their quadrature counterparts, to be interpreted as Betti numbers arising in the algebraic development. Equivalently the formulae can be derived directly from the Hilbert series of a monomial ideal associated with the base grid. The results are confirmed by computations using algorithms for Betti numbers and Hilbert series available via symbolic computation.

On sub-polynomial lower error bounds for strong approximation of SDEs

Larisa Yaroslavtseva
University of Passau

Abstract

We consider the problem of strong approximation of the solution of a stochastic differential equation (SDE) at the final time based on finitely many evaluations of the driving Brownian motion W. While the majority of results for this problem deals with equations that have globally Lipschitz continuous coefficients, such assumptions are typically not met for real world applications. In recent years a number of positive results for this problem has been established under substantially weaker assumptions on the coefficients such as global monotonicity conditions: new types of algorithms have been constructed that are easy to implement and still achieve a polynomial rate of convergence under these weaker assumptions.

In our talk we present negative results for this problem. First we show that there exist SDEs with bounded smooth coefficients such that their solutions can not be approximated by means of any kind of adaptive method with a polynomial rate of convergence. Even worse, we show that for any sequence $(a_n)_{n\in\mathbb{N}} \subset (0,\infty)$, which may converge to zero arbitrarily slowly, there exists an SDE with bounded smooth coefficients such that no approximation method based on n adaptively chosen evaluations of W on average can achieve a smaller absolute mean error than the given number a_n.

While the diffusion coefficients of these pathological SDEs are globally Lipschitz continuous, the first order partial derivatives of the drift coefficients are, essentially, of exponential growth. In the second part of the talk we show that sub-polynomial rates of convergence may happen even when the first order partial derivatives of the coefficients have at most polynomial growth, which is one of the typical assumptions in the literature on numerical approximation of SDEs with globally monotone coefficients.

The talk is based on joint work with Arnulf Jentzen (ETH Zurich) and Thomas Müller-Gronbach (University of Passau).