Abstracts

Contents

Updated: June 19, 2024

Elena Celledoni
Norwegian University of Science and Technology, Norway

Deep learning from the point of view of numerical analysis

Deep learning neural networks have recently been interpreted as discretisations of an optimal control problem subject to an ordinary differential equation constraint. A large amount of progress made in deep learning has been based on heuristic explorations, but there is a growing effort to mathematically understand the structure in existing deep learning methods and to design new approaches preserving (geometric) structure in neural networks. The (discrete) optimal control point of view to neural networks offers an interpretation of deep learning from a numerical analysis perspective and opens the way to mathematical insight [10, 9, 2].

We discuss a number of interesting directions of current and future research in structure preserving deep learning [3]. Some deep neural networks can be designed to have desirable properties such as invertibility and group equivariance or can be adapted to problems of manifold value data. Equivariant neural networks are effective in reducing the amount of data for solving certain imaging problems [4].

We show how classical results of stability of ODEs are useful to construct contractive neural networks architectures. Thus, neural networks can be designed with guaranteed stability properties. This can be used to ensure robustness against adversarial attacks and to obtain converging "Plug-and-Play" algorithms for inverse problems in imaging [3, 7, 12]. We consider extensions of these ideas to the manifold valued case and we discuss B-stability on manifolds [1].

We also consider applications of deep learning to mechanical systems, for learning Hamiltonians on manifolds and from noisy data [6, 11] and for learn-1ing PDE solutions [8]. We show how similar ideas can be used to compute optimal parametrisations in shape analysis [5].

- Lecture 1: introduction, deep learning as optimal control, dynamical systems and deep neural networks. Equivariant neural networks.

- Lecture 2: Adversarial attacks, stability of ODEs and applications to 1-Lipschitz networks and converging "Plug-and-Play" algorithms for imaging. B-stability on manifolds and applications.

- Lecture 3: Deep learning of diffeomorphisms for optimal shape reparametrization. Applications of deep learning to mechanical systems.

- Lecture 4: Learning Hamiltonians on manifolds, from noisy data and learning PDEs form pixel data.

References

[1] M Arnold, E Celledoni, E Çokaj, B Owren, D Tumiotto, (2024) *B-stability of numerical integrators on Riemannian manifolds*, Journal of Computational Dynamics.

[2] M. Benning, E. Celledoni, M. J. Ehrhardt, B. Owren, and C. B. Schönlieb, (2019) *Deep learning as optimal control problems: models and numericalmethods*. Journal of Computational Dynamics, 6(2):171–198, 2019.

[3] E. Celledoni, M. J. Ehrhardt, C. Etmann, R. I. McLachlan, B. Owren, C. B. Schönlieb, F. Sherry, (2021) *Structure preserving deep learning.* European journal of applied mathematics.

[4] E. Celledoni, M. J. Ehrhardt, C. Etmann, B. Owren, C. B. Schönlieb, F. Sherry, (2021) *Equivariant neural networks for inverse problems*. Inverse Problems.

[5] E. Celledoni, H. Glöckner, J. Riseth, A. Schmeding, (2023) *Deep neural networks on diffeomorphism groups for optimal shape reparameterization*, BIT Numerical Mathematics, 63 (4), 1-38.

[6] E. Celledoni, A. Leone, D. Murari, and B. Owren, (2023) *Learning Hamiltonians of constrained mechanical systems*, J. Comput. Appl. Math. 417, Paper No. 114608, 12 pp.

[7] E. Celledoni, D. Murari, B. Owren, C.-B. Schönlieb, and F. Sherry, (2022) *Dynamical systems' based neural networks*, SISC, 2023.

[8] E Celledoni, J Jackaman, D Murari, B Owren (2024) *Predictions based on pixel data: Insight from PDEs and finite differences*, arXiv preprint arXiv:2305.00723

[9] W. E, (2017) *A Proposal on Machine Learning via Dynamical Systems*, Commun. Math. Stat. 5, 1–11.

[10] E. Haber and L. Ruthotto, (2017) *Stable architectures for Deep Neural Networks*, Inverse Problems 34 (1).

[11] E. Celledoni, S. Eidnes, H. Noren Myhr, (2023) *Learning Dynamical Systems from Noisy Data with Inverse-Explicit Integrators*, arXiv:2306.03548

[12] F. Sherry, (2024) E. Celledoni, M. J. Ehrhardt, D. Murari, B. Owren, and C.-B. Schönlieb, *Designing Stable Neural Networks using Convex Analysis and ODEs*, Physica D: Nonlinear Phenomena 463.

Back to Contents Page →

Qiang Du
Columbia University, USA

Nonlocal modeling, analysis and computation

The world is becoming more and more nonlocal. Nonlocal models expressed as integral equations or integral-differential equations can account for nonlocal interactions and take on more general forms than discrete models, classical/local continuum models (PDEs), and fractional differential equations. Nonlocal models provide effective bridges between the discrete, local, or fractional PDE counterparts. By allowing for solutions with possibly more singular and anomalous behavior, they are well-suited for simulations of complex processes and multiscale phenomena.

Studies on Nonlocal continuum models have gained popularity in recent years with applications ranging from nonlocal diffusion and mechanics, traffic flows, pattern formation, geometric analysis of big data, and deep learning. In our lectures, we plan to briefly introduce nonlocal continuum models, their mathematical foundation and numerical discretization. The lectures are designed to contain balanced discussions on theoretical and practical issues, including illustrations of simple and motivational examples and descriptions of the state-of-art and open questions. The lectures intend to offer applied mathematicians, computational scientists, young researchers, and graduate students in many application areas a glimpse into nonlocal models' broad applicability and rich mathematics of representations.

Back to Contents Page →

Shi Jin
Shanghai Jiao Tong University, China

Quantum Computation of partial differential equations and related problems

Quantum computers have the potential to gain algebraic and even up to exponential speed up compared with its classical counterparts, and can lead to technology revolution in the 21st century. Despite its success in many areas, developing quantum algorithms for solving scientific and engineering problems remains at very early stage and many issues remain to be resolved and new mathematical thinkings are needed.

This mini-course will concentrate on quantum algorithms for solving ordinary and partial differential equations and related problems. Since quantum computers are designed based on quantum mechanics principle, they are most suitable to solve the Schrodinger equation, and linear PDEs (and ODEs) evolved by unitary operators. Other problems need to be reformulated into a form of discrete or continuous Schrodinger equation, or more specifically, into the form of unitary matrices of operators operating on quantum states.

In this mini-course, we will start with the basic knowledge about quantum computing and Schrodinger equation, and its numerical approximations. We then introduce some quantum algorithms for linear ODEs and PDEs, and related linear algebra problems.

Back to Contents Page →

Alexander Ostermann
University of Innsbruck, Austria

Time integration strategies for PDEs

The numerical integration of partial differential equations is often performed in two steps: the spatial derivatives are discretized by an appropriate numerical method (finite differences, finite elements, finite volumes, spectral methods, ...) and the resulting (non)linear system of stiff ordinary differential equations is then integrated in time. Classical time integration methods include implicit Runge-Kutta methods, implicit multistep methods, linearly implicit methods, IMEX methods, and many others. In this series of four lectures, we will focus on two widely used time integration schemes: exponential integrators and splitting methods. We will discuss the basic ideas of their construction, their numerical analysis, and their practical efficiency.

1. Exponential integrators: basics and limits

   Exponential integrators are effective time integration schemes for evolution equations and their spatial discretizations. They solve the linear part exactly and discretize the nonlinearity explicitly, resulting in high accuracy and stability when the nonlinearity is small. Exponential integrators are widely used for stiff problems arising from semilinear parabolic equations and highly oscillatory problems arising from wave or dispersive equations such as Schrödinger-type equations. The idea of exponential integrators can be traced back to the late 1950s. Their development has proceeded in several steps and has always been closely related to the possibility of efficiently computing the action of the exponential and related matrix functions. This lecture will focus on the following topics: construction, strengths and pitfalls of exponential integrators; the zoo of different methods; basic error analysis.

2. Accelerating exponential integrators

   Exponential integrators require computing the action of certain matrix functions (such as exponential and trigonometric functions) on vectors. This task is by no means independent of the chosen approximation for the vector field. Fast computations often require a particular form of the matrix, which may conflict with local linearization, often used to control the Lipschitz constant of the nonlinearity. For small problems, matrix functions are often computed explicitly, but for large problems, iterative methods such as Krylov subspace methods or Lagrange interpolation at Leja points are used. When these operations are computed efficiently, exponential integrators perform well. In important situations, acceleration techniques can be used to improve performance on modern HPC systems. This talk introduces two recent approaches: μ-mode integrators for evolution equations in

Kronecker form and accelerated methods using simplified linearization. A general strategy for selecting the appropriate integrator will also be discussed.

3. Splitting methods: basics, limits, applications

   The basic idea behind splitting methods is to divide the vector field into disjoint components, integrate them separately, and combine the results after a time step. This is a simple and often efficient procedure. However, splitting methods require careful handling of non-trivial boundary conditions. Adaptations to the integrators are necessary to address these issues. The lecture will focus on the following topics: construction and numerical analysis of splitting methods; order reduction due to non-trivial boundary conditions; an efficient combination of splitting with exponential integration in a sonic boom calculation.

4. Low regularity integration

   Standard numerical integrators such as Lie or Strang splitting and exponential integrators experience order reduction when applied to semilinear dispersive problems with non-smooth initial data. To mitigate these challenges, a recent development introduces a new class of integrators known as low-regularity integrators. These integrators use the variation-of-constants formula and employ resonance-based approximations in Fourier space, demonstrating improved convergence rates at low regularity. However, the estimation of nonlinear terms in the global error still relies on classical bilinear estimates derived from the Sobolev embedding. At very low regularity, traditional error analysis in Sobolev spaces is hampered by the lack of suitable embeddings. A novel framework, inspired by Bourgain's techniques, is developed that allows the analysis of methods applicable to very low regularity initial data. This approach is illustrated for Lie splitting applied to the `good' Boussinesq equation.

Back to Contents Page →