IDENTIFICATION OF LINEAR DYNAMICAL SYSTEMS AND MACHINE LEARNING METHODS

A. Bensoussan, F. Gelir, V. Ramakrishna, D. P. C. Nguyen , B. Tran

Abstract

The topic of identification of linear dynamical systems has been at the origin of modern control theory. It created, in particular, the domain of realization theory, consisting in realizing a linear input-ouput dynamic map, by an “internal realization”, which is a dynamical system with linear observation. We approach this classical problem with the ideas of machine learning. Machine learning aims at identifying unknown functions by least square methods with a smoothing term representing a norm in the space of unknown parameters or unknown functions.

Because of nonlinearities in the loss function, this approach, although conceptually not especially new, leads to interesting issues, regarding the choice of weights and the type of gradient methods to use. The objective is to discuss some possible algorithms.

A Large Deformation Diffeomorphic Framework for Fast Brain Image Registration via Parallel Computing and Optimization

Xiaoying Tang
Southern University of Science and Technology

Abstract

Large deformation diffeomorphic metric mapping (LDDMM) is a state-of-the-art algorithm for registering manifolds of different dimensions, such as landmarks, curves, surfaces, and dense images. The computations in LDDMM, especially when registering 3D images, are very heavy given that there will be time-dependent and voxel-wise velocity vector fields and associated diffeomorphisms. In this talk, we will introduce an efficient approach for LDDMM for brain images by utilizing GPU-based parallel computing and a mixture automatic step size estimation method for gradient descent. We systematically evaluate the proposed approach in terms of two matching cost functions, including the Sum of Squared Differences and the Cross-Correlation. Comprehensive experimental results conducted on two brain MRI datasets will be presented.

Learning Geometry for Processing Image Data

Stacey Elizabeth Levine
Duquesne University

Abstract

In recent work we developed several frameworks for image denoising that attempt to recover an image from denoised data that encodes local geometry, including the level line curvature and components in a geometrically motivated moving frame. These nonlinear transformations satisfy nice properties that provide justification for these frameworks, and the approaches are successful in practice. Still, one challenge in working with this data is that the behavior of mathematically sound mechanisms developed for handling natural image data do not readily carry over, as this data can be quite ill behaved. To mitigate this problem, in this work we use a structured convolutional neural network to learn both the geometric data from noisy observations and their corresponding regularizers. Preliminary analyses and experiments demonstrate the benefits of this approach, and suggest that the learned regularizers have the potential to feed into mathematically sound variational and PDE based approaches..

Order-p Means, M-Smoothers, and PDEs

Martin Welk
UMIT – Private University for Health Sciences,
Biomedical Informatics and Technology
Hall/Tyrol, Austria

Abstract

Order-p means are a generalisation of the classical mean, median and similar filters based on minimising a sum of p-th powers of distances in data space. They can be employed in local image filters that belong to the class of M-smoothers. In the last decades some interest has gone to order-p means with exponents less than one, which minimise non-convex objective functions, and have been associated with mode filters.

Space-continuous versions of M-smoothers can naturally be related to PDE-based image evolutions. In a suitable limit sense, mean filtering approximates the diffusion PDE, whereas median filtering approximates mean curvature motion. These results have been extended in several directions, including adaptive median filters and multivariate generalisations of medians.

In this talk, recent work in collaboration with J. Weickert on the PDE limit of the entire M-smoother family based on order-p means will be presented. A common misinterpretation about the relation between order-p mean filtering and mode filtering will be corrected in this context. This motivates at the same time an extension of the order-p class to negative exponents p which were not considered before. The resulting PDEs in one, two and three dimensions form a consistent framework overarching also some other interesting image sharpening PDEs from literature. Efficient numerical schemes for these PDEs will be discussed.

Processing High Speed Films to Extract Information from the Scene

Ola Marius Lysaker
University of South-Eastern Norway

Abstract

A framework based on high speed films and image processing for extracting physical characteristics (i.e. size and velocity distributions) in a variety of lab experiments will be presented. In the first part of the talk, the problem of estimating the size and velocity distributions for droplets in fire water sprays is discussed. This challenge is formulated as an optimization problem and solved with a Hungarian algorithm approach. Measurements of the droplet flow are obtained through a sequence of images. The size of the droplets (diameter) can be as small as 0.01 cm and the images are highly corrupted by noise.

In the second part of the talk, we discuss how neural networks and optimal transport can be applied to track the position of shock wave fronts in high speed films. More specifically, by tracking the shock wave position and its geometrical shape, characteristics like velocities and shock angles are revealed. The shock wave fronts can travel at speeds up to 4500 mph and are captured at 500 000 fps. An introduction to shock waves and how the image data was generated will be given.

Davis-Yin splitting algorithm for a class of nonconvex nonsmooth optimization

Fengmiao Bian and Xiaoqun Zhang
School of Mathematical Sciences and Institute of Natural Sciences
Shanghai Jiao Tong University

Abstract

In this paper, we consider a general optimization model, which contains a large class of practical models in data science. We employ the Davis-Yin splitting (DYS) algorithm to solve the resulting possibly nonconvex nonsmmoth optimization problems and analyze its convergence. We show that if the step-size is chosen less than a computable threshold and the sequence thus generated by DYS is bounded, then the cluster point of the sequence gives a stationary point of the nonconvex nonsmooth optimization problem. We achieve this by revealing that the sequence is decreasing along a new energy function associated with the DYS method. Furthermore, we establish the global convergence of the whole sequence under an additional assumption that this energy function is a Kurdyka-Łojasiewicz function. Finally, some numerical experiments are conducted to compare the DYS algorithm with some classical efficient algorithms for compressed sensing and low rank matrix completion. The numerical results indicate that DYS method outperforms the existing methods for these specific applications.

Image denoising via a novel deep neural network model

Wei Zhu
University of Alabama, Tuscaloosa

Abstract

In this talk, we will discuss a novel deep neural network model for image denoising. This model is based on conventional variational models and optimization techniques. Numerical experiments will be presented to show the performance of the proposed model.

Vision Models for Emerging Technologies and Their Impact on Computer Vision

Marcelo Bertalmío Barate
Pompeu Fabra University

Abstract

To enhance the overall viewing experience (for cinema, TV, games, AR/VR) the media industry is continuously striving to improve image quality. Currently the emphasis is on High Dynamic Range (HDR) and Wide Colour Gamut (WCG) technologies, which yield images with greater contrast and more vivid colours. The uptake of these technologies, however, has been hampered by the significant challenge of understanding the science behind visual perception. This talk provides an insight into the science and methods for HDR and WCG and discusses the impact on computer vision research of the limitations of current vision models. In particular, given that imaging techniques based on vision models are the ones that perform best for HDR and WCG imaging and a number of other applications, that the performance of these methods is still far below what cinema professionals can achieve, and that vision models are lacking as most key problems in visual perception remain open, we propose that rather than be improved or revisited, a change of paradigm seems to be needed for vision models, moving away from a L+NL framework. This could clearly have a really wide impact in computer vision, as the L+NL formulation is the cornerstone of artificial neural networks.

Computational approaches for coherence retrieval and beyond

Chenglong Bao (Tsinghua University)

Abstract

Coherence retrieval is an important problem in the study of light propagation and dynamics. In this talk, we firstly show the traditional measurement method results in the coherence loss due to ignoring the pixel contents. Moreover, a trance regularization has been proposed for overcoming the noise effects and an adaptive APG algorithm has been proposed for solving the resulting convex problems. The proposed numerical scheme has also been extended for solving non-convex problems such as computing phase field crystal models.

Fast Huygens Sweeping Methods for Schrodinger Equations

Shingyu Leung
Hong Kong University of Science and Technology

Abstract

In the first part of the talk, we discuss fast Huygens sweeping methods for Schrodinger equations in the semi-classical regime by incorporating short-time Wentzel-Kramers-Brillouin-Jeffreys (WKBJ) propagators into Huygens' principle. Even though the WKBJ solution is valid only for a short time period due to the occurrence of caustics, Huygens' principle allows us to construct the global-in-time semi-classical solution. Then, we will also present a more recent algorithm for simulating the multi-color optical self-focusing phenomena in nematic liquid crystals. This is a joint work with Wingfai Kwan, Jianliang Qian, Susana Serna and Xiaoping Wang.

Deep Neural Network Approach to Option Pricing in High Dimensions

Wing Lok Wan
University of Waterloo

Abstract

We propose a deep neural network framework for computing prices and deltas of American options in high dimensions. The architecture of the framework is a sequence of neural networks, where each network learns the difference of the price functions betweenadjacent time steps. We introduce the least squares residual of the associated backward stochastic differential equation as the loss function for training. Our proposed framework yields prices and deltas on the entire space time, not only at t = 0. The computationalcost of the proposed approach is quadratic in dimension, which addresses the curse of dimensionality issue that state-of-the-art approaches suffer. Our numerical simulations demonstrate these contributions, and show that the proposed neural network frameworkoutperforms state-of-the-art approaches in high dimensions.

On the Implementation of ADMM: From LASSO to Distributed Parabolic Optimal Control Problems

Xiaoming Yuan
The University of Hong Kong

Abstract

We discuss how to apply the well-known alternating direction method of multipliers (ADMM) to distributed optimal control problems with linear parabolic equation constraints. At each iteration, the main computation is for solving an unconstrained parabolic optimal control problem. Because of the high dimensionality after full discretization, it is more practical to just solve this subproblem inexactly (e.g., iteratively by the conjugate gradient method) up to medium accuracy, rather than exactly or high accuracy. Then, it becomes important to design such an appropriate inexactness criterion for these inner iterations, that can guarantee both the overall convergence (with rigorous proof) and satisfactory numerical performance (avoiding unnecessarily high accuracy for subproblems). We discuss these numerical issues, starting from a previous work on the ADMM for large-scale LASSO problems arising in statistical learning.

Trajectory of Alternating Direction Method of Multipliers and Adaptive Acceleration

Jingwei Liang
University of Cambridge

Abstract

The alternating direction method of multipliers (ADMM) is one of the most widely used first-order optimisation methods in the literature owing to its simplicity, flexibility and efficiency. Over the years, numerous efforts are made to improve the performance of the method, such as the inertial technique. In this talk, I will first discuss the limitation of inertial technique when applied to ADMM. Motivated by the failure of inertial ADMM, a framework on studying the geometric property of ADMM is provided which focuses on the trajectory of the generated sequence. Consequently, an adaptive acceleration scheme for ADMM is proposed based on the trajectory. Numerical experiments on problems arising from image processing, statistics and machine learning demonstrate the advantages of the proposed acceleration approach.

Mathematical Theory of Deep Convolutional Neural Networks

DING-XUAN ZHOU

School of Data Science and Department of Mathematics
City University of Hong Kong
Email: mazhou@cityu.edu.hk

Deep learning has been widely applied and brought breakthroughs in speech recognition, computer vision, and many other domains. The involved deep neural network architectures and computational issues have been well studied in machine learning. But there lacks a theoretical foundation for understanding the modelling, approximation or generalization ability of deep learning models with network architectures such as deep convolutional neural networks (CNNs) with convolutional structures. The convolutional architecture gives essential differences between the deep CNNs and fully-connected deep neural networks, and the classical theory for fully-connected networks developed around 30 years ago does not apply. This talk describes a mathematical theory of deep CNNs associated with the rectified linear unit (ReLU) activation function. In particular, we give the first proof for the universality of deep CNNs, meaning that a deep CNN can be used to approximate any continuous function to an arbitrary accuracy when the depth of the neural network is large enough. We also give explicit rates of approximation, and show that the approximation ability of deep CNNs is at least as good as that of fully-connected multi-layer neural networks. Our quantitative estimate, given tightly in terms of the number of free parameters to be computed, verifies the efficiency of deep CNNs in dealing with large dimensional data.

Structure preserving schemes for complex nonlinear systems

Jie Shen
Purdue University and Xiamen University

Abstract

Many complex nonlinear systems have intrinsic structures such as energy dissipation or conservation, and/or positivity/maximum principle preserving. It is desirable, sometimes necessary, to preserve these structures in a numerical scheme.

I will first present the scalar auxiliary variable (SAV) approach to deal with nonlinear terms in a large class of dissipative/conservative systems. It leads to linear and unconditionally energy stable schemes which only require solving decoupled linear equations with constant coefficients. Hence, these schemes are extremely efficient and very accurate when combined with higher-order BDF or diagonally implicit Runge-Kutta schemes. As an specific application, I shall consider a total fractional-order variational model for super-resolution. I shall also discuss its potential application for general optimization problems.

Time permitting, I will also present a strategy to construct efficient energy stable and positivity preserving schemes for certain nonlinear evolution systems, such as the Poisson-Nernst-Planck (PNP) equation and Keller-Segel equation, whose solutions remain tobe positive.

Graph models for high dimensional data and deep neural networks for images with medical applications

Xue-Cheng Tai

Department of Mathematics, Hong Kong Baptist University

Abstract

In this talk, I will present some new research work in several directions. First, we will show some graph models with applications to high dimensional data clustering. Especially, we show how to get fast algorithms using min-cut and max-flow algorithms. Moreover, we add a regional force to our model which has demonstrated to give superior accuracy for many applications. In the second part, we show that the well-know modularity maximization algorithm is in fact is volume balancing model. Using total variation on graphs, we show that we can turn the modularity maximization into a minimization problem with volume balancing property with a convex energy functional. This is a new observation and also gives some new ways to solve the modularity minimization problems.

The last part is devoted to study of deep neural networks. We propose some special techniques to add spatial regularization effects to popular deep neural networks. We use numerical experiments to show that the regularized DNN always has smooth boundary when used for image segmentation and similar classification problems. We want to emphasis that our spatial regularization effect is naturally integrated into existing deep neural networks and it only require minimal algorithmic modifications to existing neural networks. It offers very effective stability and smoothing effects into commonly used neural networks.

This talk is based join work with different collaborators.