MATRIX RESEARCH PROGRAM

SPLITTING ALGORITHMS – ADVANCES, CHALLENGES, AND OPPORTUNITIES

10 - 21 FEBRUARY 2025

CRESWICK, AUSTRALIA

https://www.matrix-inst.org.au/

https://bsky.app/profile/matrix-inst.bsky.social

https://www.linkedin.com/in/matrix-australia

Contents

Acknowledgements

We especially extend our gratitude to the tireless staff at MATRIX Institute for all of the work they have done to make our workshop possible. We would also like to acknowledge our generous sponsors.

- MATRIX Institute
- Simons Foundation
- Australian Mathematical Sciences Institute
- Australian Mathematical Society (AustMS)
- The AustMS special interest group *Mathematics of Computation and Optimisation* (MoCaO)

Invited Lecture Series

Selected Topics for Nonsmooth Optimization

Week 1 M-W

Guoyin Li
UNSW

Nonsmooth optimization problems are ubiquitous and highly challenging. In these lectures, we plan to provide a gentle introduction to some basic tools and methods for solving nonsmooth optimization problems, and to discuss a few recent developments of nonsmooth optimization. The outline of the lectures is as follows:

I Nonsmooth optimization– Brief introduction and motivation examples (1h)

II Nonsmooth optimization– Brief introduction and motivation examples (1h)

III Basic methods for solving nonsmooth optimation (1h)

- Gradient descent methods;
- Proximal gradient methods and their variants.

IV Some recent developments on nonsmooth optimization: selective topics (2h)

- Nonsmooth fractional optimization;
- The Kurdyka-Łojasiewicz inequality;
- Projection-free methods.

MATRIX

Week 1 W-F

From No Solution towards Resolution: A Nonexpansive Mapping Toolkit

Heinz Bauschke

University of British Columbia, Okanagan

Lecture 1: Classical Theory
We'll introduce the foundational concepts of nonexpansive mappings in Hilbert spaces. Our focus will include the properties of fixed points, weak convergence, and the demiclosedness principle. We'll discuss key theorems such as the Browder-Göhde-Kirk Theorem on fixed point existence and the Krasnoselskii-Mann Theorem on the convergence behavior of iterative sequences. We'll also explore Fejér monotone sequences.

Lecture 2: Maximally Monotone Operators, Pazy, Browder-Petryshyn, and Edelstein
In this lecture, we delve into the theory of maximally monotone operators in Hilbert spaces. We will examine concepts like monotonicity, the minimal displacement vector, and how nonexpansive mappings behave when they lack fixed points. Key results like Pazy's Lemma and the Browder-Petryshyn Theorem will be covered, alongside our look at Edelstein's affine isometry.

Lecture 3: Baillon-Bruck-Reich and the Minimal Displacement Vector
Join us as we investigate the behavior of averaged nonexpansive mappings, looking at their iterates and connections to monotone operators, firmly nonexpansive mappings, and nonexpansive mappings. We'll introduce significant results like the Baillon-Bruck-Reich Theorem and discuss the critical role of 3* monotonicity in operator theory. We'll also analyze how the minimal displacement vector operates under compositions and convex combinations of mappings.

Lecture 4: Cyclic Projections
We'll explore the method of cyclic projections for multiple convex sets, focusing on the Geometry Conjecture to understand the fixed points of these projections. We'll cover the behavior of alternating projections for two sets, introduce Attouch-Théra duality to extend our understanding to more sets, and resolve the Geometry Conjecture. Special cases and open problems in higher dimensions will also be discussed.

Lecture 5: Douglas-Rachford
In this lecture, we examine the Douglas-Rachford algorithm, used for solving problems with the sum of two maximally monotone operators, especially when solutions might not exist. We'll extend our discussion to include how the algorithm behaves in infeasible scenarios by introducing the concept of the normal problem. We'll also explore the Chambolle-Pock algorithm for convex optimization problems and touch upon the stability of various splitting algorithms.

Parameter-tuning free algorithm for nonconvex optimization

Week 2 M-F

Akiko Takeda

University of Tokyo/ RIKEN Center for Advanced Intelligence Project

The last decade has seen increasing interest in large-scale nonconvex optimization because of the great success of machine learning models, such as neural networks. To solve such problems, first-order optimization methods that use only function values and gradients have attracted attention. In this series of talks, after showing the theoretical guarantees for the standard gradient descent method, we will introduce some gradient-based algorithms that we have recently proposed for nonconvex optimization. First, we will present accelerated gradient methods with better iteration complexity than those of existing first-order methods. After that, for least-squares problems, we will introduce specialized algorithms that exploit the least-squares or composite structure of the objective function and are based on the Levenberg–Marquardt (LM) method. These algorithms have improved guarantees of the worst-case complexity and perform well in practice.

Week 2 M-F

Bridging continuous and discrete time models for convex optimization and monotone inclusions

Radu Ioan Boţ

University of Vienna

In this course, we will explore the interplay between continuous time and discrete time models in solving convex optimization problems and monotone inclusions. Our focus will span a broad range of problems, from unconstrained convex optimization to monotone equations, fixed-point problems, and structured monotone inclusions, as well as convex optimization problems with linear equality and cone constraints. A central theme of the course will be the striking parallels between the asymptotic behavior of continuous models and the convergence analysis of discrete algorithms. We will deal with the issues relating to the convergence of the iterates and the derivation of the accelerated rates of convergence. Additionally, we will examine strategies to enhance convergence properties, including the transition from weak to strong convergence and the development of methods to achieve faster convergence behavior. Through this lens, we aim to provide a unified understanding of both theoretical foundations and practical methodologies for tackling these problems.

Contributed Talks

Progressive Hedging for Stochastic Integer Programming - Advances in Theory and Practise

Feb 10
13:30

Andrew Eberhard
RMIT University

Motivated by recent literature demonstrating the surprising effectiveness of the heuristic implementations of progressive hedging (PH) to stochastic mixed-integer programming (SMIP) problems, we provide theoretical support for the inclusion of integer variables, filling this gap between theory and practice. We provide greater insight into the following observed phenomena of PH as applied to SMIP where the intended optimal is found or at least feasible convergence is observed. We provide an analysis of a modified PH algorithm from a different viewpoint, drawing on the interleaving of (split) proximal point methods (including PH), Gauss-Seidel methods, and the utilisation of variational analysis tools. Through this analysis, we provide conditions for convergence to a feasible solution that is observed in practise. In terms of convergence analysis, we contribute insight into the convergence of proximal point like methods in the presence of integer variables via the introduction of the notion of persistent local minima and also contribute to an enhanced Gauss-Seidel convergence analysis that accommodates the variation of the objective function. We provide a practical implementation of a modified PH and demonstrate its convergent behavior with computational experiments in line with the provided analysis.

Duality and Smooth Approximation Approaches to ADMM Acceleration

Feb 10
14:00

Andrew Calcan and Jordan Collard
Curtin University Centre for Optimisation

We will compare and contrast different approaches for using the dynamics of the Douglas—Rachford and ADMM dynamical systems to improve algorithm performance. Techniques we will consider include the Circumcentered Reflections Method, minimizing surrogates for Lyapunov functions, and Liang and Poon's Adaptive Acceleration for ADMM (A3DMM). Examples include second order cone problems, positive semidefinite feasibility, phase retrieval, compressed audio recovery, and linear regression.

Feb 10
15:00

A unified approach to distributed operator splitting

Thang D. Truong
RMIT University

Splitting methods are a class of methods which are used to decompose complex problems into smaller solvable problems. In this work, we develop a general framework of forward-backward splitting methods for solving the monotone inclusion problems involving maximally monotone operators and cocoercive or Lipschitz operators. In the literature, other attempts seem complicated and have some limitations. Our proposed frameworks, under some mild assumptions with matrix representation not only cover several important algorithms but also extend to new algorithms which will give more flexibility when applying to different problems. By choosing suitable coefficient matrices, the obtained algorithms can be implemented in distributed decentralized manner.

Feb 10
15:30

Proximal methods for nonsmooth optimization with constraints

Alberto De Marchi
University of the Bundeswehr Munich

Focusing on minimization problems with structured objective function and smooth constraints, we present some numerical methods that combine proximal algorithms with sequential unconstrained minimization techniques. Penalty and barrier schemes (with their combinations and variants) offer a diverse and flexible playground to explore. Working in the fully nonconvex setting, convergence results on par with nonlinear programming are established; stronger guarantees are often available for the same algorithms under additional conditions on the problem data. Illustrative numerical examples will be presented from a variety of application areas.

Feb 11
13:30

Faces of homogeneous cones

Bruno Lourenço
Institute of Statistical Mathematics

A convex cone is said to be homogeneous, if its group of automorphisms acts transitively on its relative interior. In this talk, we will present our recent results on the facial structure of homogeneous cones and discuss applications to sparse PSD cones involving homogeneous chordality and the corresponding dual cone of PSD completable matrices. This is a joint work with Masaru Ito (Nihon University) and João Gouveia (University of Coimbra)

MATRIX

Finite Convergence of Circumcentered-Reflection Method on Closed Polyhedral Cones in Euclidean Spaces

Feb 11
14:00

Hongzhi Liao
UNSW Sydney

The Circumcentered Reflection Method (CRM) is a recently developed projection method for solving convex feasibility problems. It offers preferable convergence properties compared to classic methods such as the Douglas-Rachford and the alternating projections method. In this study, our first main theorem establishes that CRM can identify a feasible point in the intersection of two closed convex cones in $\mathbb{R}^2$ from any starting point in the Euclidean plane. We then apply this theorem to intersections of two polyhedral sets in $\mathbb{R}^2$ and two wedge-like sets in $\mathbb{R}^n$, proving that CRM converges to a point in the intersection from any initial position finitely. Additionally, we introduce a modified technique based on CRM, called the Sphere-Centered Reflection Method. With the help of this technique, we demonstrate that CRM can locate a feasible point in finitely many iterations in the intersection of two proper polyhedral cones in $\mathbb{R}^3$ when the initial point lies in a subset of the complement of the intersection's polar cone. Lastly, we provide an example illustrating that finite convergence may fail for the intersection of two proper polyhedral cones in $\mathbb{R}^3$ if the initial guess is outside the designated set.

Optimal control duality and the Douglas–Rachford algorithm

Feb 11
15:00

Bethany Caldwell
University of South Australia

The Douglas–Rachford (DR) algorithm is a popular and well-studied algorithm but despite its prevalence there have been few studies applying this to infinite dimensional problems. We apply this algorithm to an infinite dimensional optimization problem, specifically, the dual of a weighted minimum energy linear-quadratic control problem with control constraints. We derive the fixed point of the DR operator as applied to this problem and using this expression for the fixed point we construct an optimality check to verify our numerical solutions. We propose an algorithm for our optimal control problem that gives both the primal and dual sequences and apply this to two example optimal control problems.

MATRIX

Feb 11
15:30

Proximal splitting methods through the lenses of Moreau-type envelopes

Felipe Atenas
University of Melbourne

The proximal point algorithm for nonsmooth optimization can be interpreted as the gradient method applied to a regularized merit function, the Moreau envelope. A similar interpretation can be given to proximal-type splitting methods that solve structured optimization problems, using envelope functions tailored to the structure of such problems and the specific method. In this talk, using merit functions resembling the Moreau envelope, we will discuss how to derive theoretical and algorithmic properties for two common two-operator splitting methods, the forward-backward and the Douglas-Rachford methods. We will explore the smoothness and epi-approximation properties of the envelopes and how they explain the convergence of these methods in nonconvex settings.

Feb 17
13:30

Some novel convergence rates for common fixed point problems

Tianxiang Liu
Tokyo Institute of Technology

In this talk, we will present some novel and concrete convergence rates for common fixed point problems involving Karamata regular operators. These operators include but can be beyond Holderian ones, in particular when the problem involves the exponentials and logarithms. With the aid of Karamata theory, the rates obtained then include those which sit between linear and sublinear rates, for example, expressed by the Lambert W function, in a case when the Douglas-Rachford algorithm is applied to nonsemialgebraic sets. We will also connect the existence of such operators to o-minimal geometry.

Feb 17
14:00

Fixed-time convergent second-order time-varying dynamical systems

Lien Nguyen
RMIT University

Addressing optimization problems using the finite-time/fixed-time stability properties of dynamical systems has received increasing interest. In this work, we propose new fixed-time second-order time-varying dynamical systems. We derive a sufficient condition for the existence and uniqueness of solutions to a general ordinary differential equation (ODE) and present a rigorous proof establishing the well-posedness of our proposed algorithms, addressing a significant gap in the literature. We then employ the proposed dynamical systems to solve nonsmooth additive composite optimization problems.

Convergence of the Douglas-Rachford Method for Conic Feasibility Problems Without Slater's Condition

Feb 17
15:00

Liam Walsh

Curtin University

Conic feasibility problems involve finding a point at the intersection of an affine subspace and a closed convex cone. The Douglas-Rachford method is a well-established splitting algorithm for solving such problems, with guaranteed R-linear convergence under the assumption of Slater's condition. In this talk, we discuss the convergence behaviour of the Douglas-Rachford method when Slater's condition is not satisfied. Our framework includes the cases of p-cones and the exponential cone, wherein we exploit recent error bound results.

Douglas-Rachford algorithm for nonmonotone multioperator inclusion problems

Feb 17
15:30

Jan Harold Alcantara

University of Tokyo/ RIKEN Center for Advanced Intelligence Project

The Douglas-Rachford algorithm is a classic splitting method for finding a zero of the sum of two maximal monotone operators. This framework has been extended to cases involving a weakly monotone and a strongly monotone operator, provided a sufficiently small step size is used. In this work, we apply the Douglas-Rachford algorithm to the multioperator inclusion problem, reformulated as a two-operator inclusion in a product space. We show that when the operators are weakly monotone and strongly monotone, the algorithm converges globally to a solution of the inclusion problem under sufficiently small step sizes. Additionally, we demonstrate applications to unconstrained multiblock optimization problems with weakly convex and strongly convex functions. For general nonconvex problems in finite-dimensional spaces involving Lipschitz continuously differentiable functions and a proper closed function, we further establish global subsequential convergence results.

Implicit bias in optimization algorithms

Feb 18
13:30

Yura Malitsky

University of Vienna

In this talk, we will explore the role of implicit bias in optimization algorithms. Implicit bias refers to the tendency of an algorithm to converge to a specific solution even in the absence of an explicit regularization term in its formulation. In other words, implicit bias emerges from how the algorithm itself interacts with the objective function it's trying to minimize. We will demonstrate this concept through a few surprising applications.

Variational properties of the abstract subdifferential operator

Feb 18
14:00

Reinier Diaz Millan
Deakin University

Abstract convexity generalises classical convexity by considering the suprema of functions taken from an arbitrarily defined set of functions. These are called the abstract linear (abstract affine) functions. The purpose of this paper is to study the abstract subdifferential. We obtain a number of results on the calculus of this subdifferential: summation and composition rules and prove that under some reasonable conditions, the subdifferential is a maximal abstract monotone operator. Another contribution of this paper is a counterexample demonstrating that the separation theorem between two abstract convex sets is generally untrue.

Heavy-ball ODE converges at rate $O(T^{-4/7})$ for nonconvex functions

Feb 18
14:30

Naoki Marumo
University of Tokyo

Several algorithms based on Nesterov's accelerated gradient descent (AGD) or Polyak's heavy-ball (HB) method can find an ε-stationary point within $O(\varepsilon^{-7/4})$ or $\tilde{O}(\varepsilon^{-7/4})$ gradient evaluations under Lipschitz continuous gradients and Hessians. These methods incorporate additional mechanisms, such as restart schemes and negative curvature exploitation, which complicate the algorithms' behavior. To explore the feasibility of a simpler algorithm, we investigate the HB differential equation, a continuous-time analogy of AGD and HB; we prove that its dynamics attain an ε-stationary point within $O(\varepsilon^{-7/4})$ time.
