VGSCO

Vienna Graduate School on Computational Optimization

Vienna Workshop on Computational Optimization

December 17-19, 2018 | University of Vienna

Dear participants

The Vienna Workshop on Computational Optimization is organized within the framework of the Vienna Graduate School on Computational Optimization (VGSCO), a doctoral program financed by the Austrian Science Fund (FWF). This program started in October 2016 and has the intention to give our PhD candidates a comprehensive training in all areas of computational optimization, ranging from continuous convex and nonconvex optimization, global optimization, combinatorial and discrete optimization, dynamic and stochastic optimization to heuristic optimization methods and algorithms in computer science.

Another goal of the education is to cover the whole scope from theory over algorithmic implementations to applications in various fields. It is in this spirit that the program of this workshop was designed. With our prominent invited speakers and with the selected speakers, this workshop presents a wide range of topics in optimization with a view on algorithmic solutions. We are convinced that this event will contribute to strengthen the role of computational optimization at the participating institutions - the University of Vienna (Department of Mathematics, Department of Statistics and OR and Department of Computer Science) as well as the University of Technology in Vienna (Algorithms and Complexity Group).

We wish you an informative and pleasent stay in Vienna.

Georg Pflug on behalf of the VGSCO - Faculty
(Immanuel Bomze, Radu Bot, Monika Henzinger, Arnold Neumaier, Günther Raidl, Hermann Schichl)

Vienna Workshop on Computational Optimization

University of Vienna, December 17-19, 2018

Venue: Sky Lounge, 12th floor, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria

Plenary speakers

- H. Bauschke, University of British Columbia
- S. P. Boyd, Stanford University
- P. Gritzmann, Technical University of Munich
- H. Raecke, Technical University of Munich
- A. Schoebel, University of Goettingen
- S. J. Wright, University of Wisconsin

Program (preliminary version)

Monday, December 17, 2018

Morning session

8:30-9:00 Registration/opening

9:00-10:00 S. Wright: Optimization in Data Analysis: Some Recent Developments

10:00-10:30 M. Avolio, A. Fuduli: Proximal support vector machine variants for binary multiple instance learning

10:30-11:00 Coffee break

11:00-11:30 S. Banert, A. Ringh, J. Adler, J. Karlsson, O. Öktem: Data-driven nonsmooth optimization

11:30 12:00 A. Astorino, A. M. Bagirov, A. Fuduli: A piecewise linear support vector machines approach

12:00-12:30 C. Bertocchi, M.-C. Corbineau, E. Chouzenoux, J.-C. Pesquet, M. Prato: A novel proximal interior point neural network approach for image recovery

Afternoon Session

14:00-15:00 A. Brieden, P. Gritzmann, F. Klemm: Diagrams and democracy: electoral district design via constrained clustering

15:00-15:30 M. Staudigl, P. Mertikopoulos, I. Bomze, W. Schachinger: On the convergence of projection free Hessian barrier gradient algorithms

15:30-16:00 Coffee break

16:00-16:30 M. Labbé, F. Plein, M. Schmidt: Bookings in the European Gas Market: Characterisation of Feasibility and Computational Complexity Results

16:30-16:50 A. Böhm, R. I. Bot: A variable smoothing algorithm for convex optimization problems using stochastic gradients

16:50-17:10 C. Geiersbach, E. Loayza, K. Welker: Shape Optimization for Interface Identification under Uncertainty

17:10-17:30 R.I. Bot, E.R. Csetnek, D.-K. Nguyen: A proximal minimization algorithm for structured nonconvex and nonsmooth problems

17:30-17:50 M. Kimiaei, A. Neumaier: Efficient global unconstrained black box optimization

Tuesday, December 18, 2018

Morning Session

9:00-10:00 S. Boyd: Convex optimization

10:00-10:30 D. Ghilli, K. Kunisch: Theory and numerical practice for optimization problems involving p-functionals, with $p \in [0, 1)$

10:30-11:00 Coffee break

11:00-11:30 F. Mannel, A. Rund: Quasi-Newton methods in structured nonsmooth optimization a hybrid approach

11:30-12:00 J. Vidal-Nunez, R. Herzog, R. Bergmann, S. Schmidt, M. Herrmann: Variational mesh denoising and surface fairing using the total variation of the normal

12:00-12:30 A. Viorel: Optimization algorithms based on operator splitting and energy-preserving numerical integrator

Afternoon Session

14:00-15:00 H. Raecke: Polylogarithmic guarantees for generalized reordering buffer management

15:00-15:30 V. Kolmogorov: Valued constraint satisfaction problems

15:30-16:00 Coffee break

16:00-16:30 C. Kümmerle, C. M. Verdun: Denoising and Completion of Structured Low-Rank Matrices via Iteratively Reweighted Least Squares

16:30-17:00 A. Masone, A. Sforza, C. Sterle, A. Ushakov, I. Vasilyev: A p-median based exact method for the large-scale optimal diversity management problem

17:00 Walk to Christmas market in AKH

Wednesday, December 19, 2018

Morning Session

9:00-10:00 H. Bauschke: On the displacement of compositions and convex combinations of nonexpansive operators

10:00-10:30 Y. Malitsky: Bilevel composite minimization problems

10:30-11:00 Coffee break

11:00-11:30 S. C. Laszlo: Convergence rates for an inertial algorithm of gradient type associated to a smooth nonconvex minimization

11:30-12:00 S. Rebegoldi, S. Bonettini, V. Ruggiero: A variable metric approach for the inexact inertial forwardbackward method

12:00-12:30 S.-M. Grad, O. Wilfer: Solving nonlinear minmax location problems with minimal time functions by means of conjugate duality and proximal point methods

Afternoon Session

14:00-15:00 A. Schoebel: New concepts in robust optimization

15:00-15:30 B. Rudloff, G.Kovacova, Z. Feinstein: Dynamic Multivariate Programming

15:30-16:00 Coffee break

16:00-16:30 D. Mohseni-Taheri, S. Nadarajah, A. Trivella: A dual reoptimization scheme for solving large-scale stochastic optimization problems

16:30-17:00 G. Bigi, S. Sagratella: Robustness and semi-infinite programming via Nash games

17:00-17:30 L. Altangerel: Gap functions for quasi-variational inequality via duality

Abstracts of the Plenary Speakers of the Vienna Workshop on Computational Optimization?

On the displacement of compositions and convex combinations of nonexpansive operators

Heinz Bauschke

Mathematics, The University of British Columbia, Kelowna, B.C., Canada.
heinz.bauschke@ubc.ca

Various operators in optimization are compositions or convex combinations of other (simpler) operators. Popular examples include operators from splitting methods such as forward-backward or Douglas–Rachford. In this talk, which is based on joint work with Walaa Moursi (Stanford and Waterloo), I will report on recent progress on understanding these operators in the (possible) absence of fixed points.

Convex Optimization

Stephen Boyd

Stanford University
boyd@stanford.edu

Convex optimization has emerged as useful tool for applications that include data analysis and model fitting, resource allocation, engineering design, network design and optimization, finance, and control and signal processing. After an overview of the mathematics, algorithms, and software frameworks for convex optimization, we turn to common themes that arise across applications, such as sparsity and relaxation. We describe recent work on real-time embedded convex optimization, in which small problems are solved repeatedly in millisecond or microsecond time frames, and large-scale distributed convex optimization, in which many solvers are coordinated to solve enormous problems.

Diagrams and Democracy: Electoral District Design via Constrained Clustering

A. Brieden[1], P. Gritzmann[2], and F. Klemm[3]

[1] *Universität der Bundeswehr, 85579 Neubiberg, Germany*
andreas.brieden@unibw.de[1]
[2−3] *Department of Mathematics, Technical University Munich*
gritzman@ma.tum.de[2], **klemm@ma.tum.de**[3]

We study the electoral district design problem where municipalities of a state have to be grouped into districts of nearly equal population while obeying certain politically motivated requirements. We develop a general framework for electoral district design that is based on the close connection of constrained geometric clustering and diagrams. The approach is computationally efficient and flexible enough to pursue various conflicting juridical demands for the shape of the districts. We demonstrate the practicability of our methodology for electoral districting in Germany (and for other applications).

References

[1] Brieden, A., P. Gritzmann; F. Klemm. Constrained clustering via diagrams: A unified theory and its application to electoral district design in European Journal of Operational Research 263, 2017, pages 18-34

Polylogarithmic Guarantees For Generalized Reordering Buffer Management

Harald Räcke

Computer Science Department, Technical University Munich
raecke@in.tum.de

In the generalized reordering buffer management problem you are given k servers that have to process a sequence of requests in a metric space. There exists a buffer that at any point in time can hold up to b requests. In each step an online algorithm has to choose a request r to be served and a server s to serve it. Then s moves to the location of r and the next request from the requst sequence takes r's place in the buffer. The goal is to minimize the total distance travelled by the servers. This problem is a natural combination of the k-server problem and the reordering buffer management problem. We show how to obtain a competitive ratio of $O((logk + loglogb)logk)$ for a uniform metric space. For $k = 1$ the lower bound on the competitive ratio is $\Omega(loglogb)$ while for$b = 1$ the lower bound is $\Omega(logk)$ from the paging problem.

References

[1] Matthias Englert, Harald Räcke, and Richard Stotz. Polylogarithmic Guarantees for Generalized Reordering Buffer Management. Submitted.

New Concepts in Robust Optimization

Anita Schöbel

Georg August University Göttingen
schoebel@math.uni-goettingen.de

Most real-world optimization problems contain parameters which are not known at the time a decision is to be made. In robust optimization one specifies the uncertainty in a scenario set and tries to hedge against the worst case. Classical robust optimization aims at finding a solution which is best in the worst-case scenario. Being a well-studied concept [1], it is known to be very conservative. This motivated researchers to introduce less conservative robustness concepts in the last decade, see [2].
In the first part of this talk, two of such less conservative robustness approaches are introduced and discussed: Light robustness and a scenario-based approach to recovery robustness. While light robustness ensures a pre-defined nominal quality, recovery robustness allows to adapt a solution if the true scenario becomes known.
The second and main part of the talk goes one step further: How to handle uncertain optimization problems in which more than one objective function is to be considered? This yields a robust multiobjective optimization problem, a class of problems only recently introduced and studied, see [3, 4]. We introduce different concepts on how to define a robust Pareto solution, discuss their relations to set-valued optimization and algorithmic approaches.

References

[1] Ben-Tal, A., El Ghaoui, L., Nemirovski, A. (2009) Robust Optimization, Princeton University Press.

[2] Goerigk, M. and Schöbel, A. (2016) Algorithm Engineering in Robust Optimization, in: Algorithm Engineering: Selected Results and Surveys, ed: L. Kliemann and P. Sanders, LNCS 9220, 245-279.

[3] Ide, J. and Schöbel, A. (2016) Robustness for uncertain multi-objective optimization: a survey and analysis, OR Spectrum 38:1, 235-271.

[4] Botte, M. and Schöbel, A. (2018) Dominance for Multi-Objective Robust Optimization Concepts, European Journal of Operational Research, to appear

Optimization in Data Analysis: Some Recent Developments

Stephen J. Wright

Computer Sciences Department and Wisconsin Institute for Discovery, University of Wisconsin-Madison
swright@cs.wisc.edu

Optimization is vital to the modern revolution in data science, and techniques from optimization have become essential in formulating and solving a wide variety of data analysis problems. In turn, data science has caused a ferment of new research activity in optimization by posing challenging new problems and new contexts. We start this talk with an overview of the many problem classes in data science in which optimization provides the key solution methodology. We then focus on several areas of growing recent interest, including the interplay between optimization and data analysis in such areas as nonconvex optimization, robust optimization, adversarial machine learning, neural networks, and matrix optimization.

Abstracts of the Participants of the Vienna Workshop on Computational Optimization

Gap Functions for Quasi-variational Inequalities via Duality

Lkhamsuren Altangerel

Faculty of Mathematics, Computer and Natural Sciences; German-Mongolian Institute for Resources and Technology
altangerel@gmit.edu.mn

This paper deals with an application of duality results from [4] which deals with minimization of a convex function over the solution set of a range inclusion problem determined by set-valued mapping to the construction of gap functions for quasi-variational inequalities. The same approach was investigated for variational inequalities and equilibrium problems in [1,2] and the study shows that we can obtain some previous results for variational inequalities as special cases (cf.[2,5]). Moreover, some applications dealing with the generalized Nash equilibrium problems and mixed variational inequalities are presented.

References

[1] Altangerel, L.; Bot, R.I.; Wanka, G. (2006) On gap functions for equilibrium problems via Fenchel d uality, Pacific Journal of Optimization 2 (3), 667–678.

[2] Altangerel, L.; Bot, R.I.; Wanka, G. (2007) On the construction of gap functions for variational inequalities via conjugate duality, Asia-Pacific Journal of Operational Research 24, No. 3, 353–371.

[3] Altangerel, L. (2018) Gap functions for quasi-variational inequalities via duality, Journal of Inequalities and Applications 1, 13.

[4] Bot, R.I; Csetnek, E.R. (2013) Conjugate duality and the control of linear discrete systems, Journal Optimization Theory and Applications 159, No. 3, 576-589.

[5] Fukushima, M. (2007) A class of gap functions for quasi-variational inequality problems, Journal of Industrial and Management Optimization 3, No.2, 165- 171.

A piecewise linear Support Vector Machines approach

Annabella Astorino[1], Adil M. Bagirov[2], and Antonio Fuduli[3]

[1] *Institute of High Performance Computing and Networking, Italian National Research Council, Rende, Italy*
annabella.astorino@icar.cnr.it
[2] *School of Science, Engineering and Information Technology, Federation University Australia, Ballarat, Victoria, Australia*
a.bagirov@federation.edu.au
[3] *Department of Mathematics and Computer Science, University of Calabria, Rende, Italy*
antonio.fuduli@unical.it

The problem of separating finite sets has many applications in applied mathematics and data mining. If the sets are linearly separable then one hyperplane provides complete separation, however, in many real-world applications, this is not the case. In most datasets, in fact, classes are disjoint but their convex hulls intersect. In this situation, the decision boundary between the classes is nonlinear and it can be approximated using piecewise linear functions [1, 2]. In this work a general approach to design piecewise linear Support Vector Machines is proposed. Using max-min linear separability the notion of a margin is defined and then it is introduced in the classification error function which can be put in a difference of convex (DC) form. An incremental approach [3] is used to design an algorithm for finding starting points to minimize the nonconvex error function. Piecewise linear classifiers for both polyhedral and max-min separabilities are introduced. Finally these classifiers are tested on some large real-world data sets and the results of numerical experiments are reported.

References

[1] Astorino, A. and Gaudioso, M. (2002). Polyhedral separability through successive LP, Journal of Optimization Theory and Applications, 112, 265-293.

[2] Bagirov, A. M. (2005). Max-min separability, Optimization Methods and Software, 20, 271-290.

[3] Bagirov, A. M., Ugon, J. and Webb, D. (2011). An efficient algorithm for the incremental construction of a piecewise linear classifier, Information Systems, 36, 782-790.

Data-driven nonsmooth optimization

Sebastian Banert[1], Axel Ringh[2], Jonas Adler[3], Johan Karlsson[4] and Ozan Öktem[5]

[1−5] *KTH Royal Institute of Technology, 100 44 Stockholm, Sweden*
banert@kth.se[1], aringh@kth.se[2], jonasadl@kth.se[3], johan.karlsson@math.kth.se[4], ozan@kth.se[5]

In this talk, we discuss methods for solving large-scale optimization problems with a possibly nonsmooth objective [1]. The key idea is to first specify a class of optimization algorithms using a generic iterative scheme involving only linear operations and applications of proximal operators. This generic scheme contains several modern primal-dual first-order solvers like the Chambolle-Pock algorithm [2] as special cases. Moreover, we show convergence to a solution for a new method which also belongs to this class. Next, we interpret the generic scheme as a recurrent neural network and use unsupervised training to learn the best set of parameters for a specific class of objective functions while imposing a fixed number of iterations. We demonstrate this principle of "learning to optimize" for the family of total variation regularization problems. In contrast to other instances of this principle, e.g. in [3], we present an approach which learns parameters only in the set of convergent schemes. As use cases, we consider optimization problems arising in tomographic reconstruction and image deconvolution.

References

[1] Banert, S., Ringh, A., Adler, J., Karlsson, J., and Öktem, O. (2018). Datadriven nonsmooth optimization. Preprint. http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-229649

[2] Chambolle, A. and Pock, T. (2011). A first-order primal-dual algorithm for convex problems with applications to imaging. Journal of Mathematical Imaging and Vision, 40(1), pp. 120–145.

[3] Gregor, K. and LeCun, Y. (2010). Learning fast approximations of sparse coding. In: Proceedings of the 27th International Conference on Machine Learning (ICML-10), pp. 399-406.

A novel proximal interior point neural network approach for image recovery

C. Bertocchi[1], M.-C. Corbineau[2], E. Chouzenoux[3], J.-C. Pesquet[4] and M. Prato[5]

[1,5] *Dipartimento di Scienze Fisiche, Informatiche e Matematiche, Università degli Studi di Modena e Reggio Emilia, Via Campi 213/b, 41125 Modena, Italy*
carla.bertocchi@unimore.it[1]**, marco.prato@unimore.it**[5]
[2,4] *Center for Visual Computing, INRIA, CentraleSupèlec, University Paris-Saclay, 9 rue Joliot Curie, 91190 Gif sur Yvette, France*
mariecaroline corbineau@hotmail.fr[2]**,**
jean-christophe.pesquet@centralesupelec.fr[4]
[3] *University Paris-East, LIGM, UMR CNRS 8049, 2 Bvd Blaise Pascal - Citè Descartes, 93162 Noisy-Le-Grand, France*
emilie.chouzenoux@centralesupelec.fr[3]

We propose a novel neural network to approach the image restoration problem. This architecture is inspired from an interior point proximal optimization algorithm, capable of imposing useful constraints on the sought solution [1]. In particular, the network is composed of proximal steps alternated with convolutional structures that are able to estimate in an automatic manner the involved parameters, such as the regularization parameter, the steplength and the barrier parameter. This is one of the advantages offered by the proposed network with respect to variational methods traditionally employed in image restoration, for which the choice of parameters is performed either empirically or with suboptimal techniques. Also numerical experiments for image deblurring/denoising show that our network trained in a supervised fashion is much faster and leads to a better restoration quality than standard optimization methods.

References

[1] Combettes P.L., Pesquet JC. (2011). Proximal splitting methods in signal processing. In: Bauschke H., Burachik R., Combettes P., Elser V., Luke D., Wolkowicz H. (eds) Fixed-Point Algorithms for Inverse Problems in Science and Engineering. Springer Optimization and Its Applications, vol 49. Springer, New York, NY

Robustness and semi-infinite programming via Nash games

Giancarlo Bigi[1] and Simone Sagratella[2]

[1] *Dipartimento di Informatica, Università di Pisa*
giancarlo.bigi@unipi.it
[2] *Dipartimento di Ingegneria Informatica, Automatica e Gestionale, Sapienza Università di Roma*
sagratella@diag.uniroma1.it

One way to deal with uncertainty in optimization problems relies on the introduction of uncertainty sets for the data. This allows considering solutions that are feasible for any realization of the data while taking into account the worst-case for the objective as well. A problem of production planning under price uncertainty is exploited to address how to formulate robust counterparts of optimization problems as semi-infinite programs (shortly SIPs), i.e., optimization problems with infinitely many constraints. In turn, SIPs share some similarities with Generalized Nash games (shortly GNEPs) that lead to meaningful connections. Indeed, SIPs can be reformulated as GNEPs with a peculiar structure under some mild assumptions. Pairing this structure with a suitable partial penalization scheme for GNEPs leads to a class of solution methods for SIPs that are based on a sequence of saddlepoint problems. Any converging algorithm for the saddlepoint problem provides the basic iterations to perform the penalty updating scheme. In particular, a projected subgradient method for nonsmooth optimization and a subgradient method for saddlepoints are adapted to our framework and the convergence of the resulting algorithms is shown. A comparison between the two algorithms is outlined as well.

References

[1] Combettes P.L., Pesquet JC. (2011). Proximal splitting methods in signal processing. In: Bauschke H., Burachik R., Combettes P., Elser V., Luke D., Wolkowicz H. (eds) Fixed-Point Algorithms for Inverse Problems in Science and Engineering. Springer Optimization and Its Applications, vol 49. Springer, New York, NY

A variable smoothing algorithm for convex optimization problems using stochastic gradients

Axel Böhm[1] and Radu Ioan Bot[2]

[1−2] *Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria*

axel.boehm@univie.ac.at[1], **radu.bot@univie.ac.at**[2]

In this work we aim to solve a structured convex optimization problem, where a non-smooth function is composed with a linear operator. When opting for full splitting schemes, usually, primal-dual type methods are employed as they are effective and also well studied. However, under the additional assumption of Lipschitz continuity of parts of the objective function we can derive novel algorithms through regularization via the Moreau envelope. Applications can be found e.g. in inverse problems which lend themselves to the application of stochastic methods by the means of gradient estimators.

Proximal Support Vector Machine variants for binary Multiple Instance Learning

Matteo Avolio[1] and Antonio Fuduli[2]

[1−2] *Dipartimento di Matematica e Informatica, Universitá della Calabria, Italy*
matteoavolio0@gmail.com[1], **antonio.fuduli@unical.it**[2]

We tackle a binary Multiple Instance Learning (MIL) problem, where the objective is to discriminate between two classes of sets of points: positive and negative. Such sets are called bags and the points inside each bag are called instances. In particular, we focus on the case with two classes of instances, where a bag is positive when it contains at least a positive instance and it is negative when all its instances are negative. A well-established instance-level SVM type model for such kind of problems has been proposed in [1] and the related optimization techniques are based on solving iteratively a finite number of quadratic programming problems of the SVM (Support Vector Machine) type, which may require long computational times. Then, in order to speed up the optimization process, we propose a different model based on the PSVM (Proximal Support Vector Machine) approach introduced in [2] by Fung and Mangasarian for the standard supervised learning. Preliminary computational results are presented on a set of MIL test problems drawn from literature. Finally some variants based on the Lagragian relaxation technique are considered.

References

[1] S. Andrews, I. Tsochantaridis, T. Hofmann, (2003). Support vector machines for multiple-instance learning. in Advances in Neural Information Processing Systems, S. Becker, S. Thrun, and K. Obermayer, eds., MIT Press, Cambridge, 2003, pp. 561–568.

[2] G. Fung, O.L. Mangasarian, (2001). Proximal support vector machine classifiers. In F. Provost and R. Srikant, editors, Proceedings KDD-2001: Knowledge Discovery and Data Mining, August 26-29, 2001, San Francisco, CA, pp. 77-86.

Shape Optimization for Interface Identification under Uncertainty

Caroline Geiersbach[1], Estefania Loayza[2] und Kathrin Welker[3]

[1] *Department of Statistics and Operations Research, University of Vienna*
caroline.geiersbach@univie.ac.at
[2] *Faculty of Mathematics, Technische Universität Chemnitz*
estefania.loayza@math.tu-chemnitz.de
[3] *Department of Mathematics, Universität Trier*
welker@uni-trier.de

In shape optimization, a shape is sought that optimizes certain response properties. Many relevant problems involve constraints in the form of a partial differential equation. Such cons- traints contain inputs, such as random forces or material properties, that might not be known exactly. In this talk, we present an inverse problem with stochastic modeling, which involves the identification of an interface between two materials based on measurements made on the boundary of the domain. The model will be cast as a stochastic optimization problem of the form

$$\min_{D \in \mathcal{D}} \mathbb{E}[J_\omega(D)]$$

where $D \subset \mathbb{R}^2$ is a shape from a set of admissible shapes $\mathcal{D}$. A hybrid approach to solving this problem numerically will be developed, which relies on optimization over an abstract shape manifold combined with stochastic approximation. Numerical experiments will demonstrate the effectiveness of the approach.

Theory and numerical practice for optimization problems involving ℓ^p-functionals, with $p \in [0, 1)$

Daria Ghilli[1] and Karl Kunisch[2]

[1] *Institute of Mathematics, University of Graz*
daria.ghilli@uni-graz.at
[2] *Institute of Mathematics, University of Graz and RICAM Institute, Linz*
karl.kunisch@uni-graz.at

Nonsmooth nonconvex optimization problems involving the ℓ^p quasi-norm, $p \in [0, 1)$, of a linear map are considered. A monotonically convergent scheme for a regularized version of the original problem is developed and necessary optimality conditions for the original problem in the form of a complementary system amenable for computation are given. Then an algorithm for solving the above mentioned necessary optimality conditions is proposed. It is based on a combination of the monotone scheme and a primal-dual active set strategy. The performance of the two schemes is studied and compared to other existing algorithms by means of a series of numerical tests in different cases, including optimal control problems, fracture mechanics and microscopy image reconstruction.

Solving nonlinear minmax location problems with minimal time functions by means of conjugate duality and proximal point methods

Sorin-Mihai Grad[1] and Oleg Wilfer[2]

[1]*Faculty of Mathematics and Computer Science, Leipzig University, Augustusplatz 10, 04109 Leipzig, Germany*
grad@math.tu-chemnitz.de
[2]*Faculty of Mathematics, Chemnitz University of Technology, D-09107 Chemnitz, Germany*
oleg.wilfer@mathematik.tu-chemnitz.de

We investigate via a conjugate duality approach general nonlinear minimax location problems formulated by means of minimal time functions, necessary and sufficient optimality conditions being delivered together with characterizations of the optimal solutions in some particular instances. A splitting proximal point method is employed in order to numerically solve such problems and their duals and we present the computational results obtained in matlab on concrete examples, comparing these with earlier similar ones from the literature.

Efficient global unconstrained black box optimization

Morteza Kimiaei[1] and Arnold Neumaier[2]

[1−2]*Fakultät für Mathematik, Universität Wien, Oskar-Morgenstern-Platz 1, A-1090 Wien, Austria*
kimiaeim83@univie.ac.at[1], **arnold.Neumaier@univie.ac.at**[2]

For the unconstrained global optimization of black box functions, this paper presents a new stochastic algorithm called VSBBO. In practice, VSBBO matches the quality of other state-of-the-art algorithms for finding, with reasonable accuracy, a global minimizer in small and large dimensions, or at least in the majority of cases a point as good as competing algorithms.

For smooth, everywhere defined functions, it is proved that, with probability arbitrarily close to 1, one finds with $O(n\epsilon^{-2})$ function evaluations a point with gradient 2-norm $\leq \epsilon$. In the smooth convex case, this number improves to $O(n\epsilon^{-1})$ and in the smooth strongly convex case to $O(n \log \epsilon^{-1})$. This matches the recent complexity results by Bergou, Gorbunov and Richtárik for reaching a slightly different goal, namely the expected gradient 2-norm $\leq \epsilon$.

References

[1] M. Kimiaei , A. Neumaier, LMBOPT - a limited memory method for bound-constrained optimization, in preprint (2018).

[2] E.H. Bergou, E. Gorbunov and P. Richtárik, Random direct search method for minimizing nonconvex, convex and strongly convex functions, Manuscript (2018).

[3] N.I.M. Gould, D. Orban, Ph.L. Toint, CUTEst: a constrained and unconstrained testing environment with safe threads for mathematical optimization. Computational Optimization and Applications, 60(3) (2015), 545-557.

[4] W. Huyer and A. Neumaier, Global optimization by multilevel coordinate search, J. Global Optimization 14 (1999), 331-355.

[5] W. Huyer and A. Neumaier, SNOBFIT-stable noisy optimization by branch and Fit, ACM Trans. Math. Software 35, Article 9 (2008).

[6] A. Neumaier, H. Fendl, H. Schilly and T. Leitner, Derivative-free unconstrained optimization based on QR factorizations, Soft Computing 15 (2011), 2287-2298.

Valued Constraint Satisfaction Problems

Vladimir Kolmogorov

Institute of Science and Technology Austria
vnk@ist.ac.at

I will consider the Valued Constraint Satisfaction Problem (VCSP), whose goal is to minimize a sum of local terms where each term comes from a fixed set of functions (called a "language") over a fixed discrete domain. I will present recent results characterizing languages that can be solved using the basic LP relaxation. This includes languages consisting of submodular functions, as well as their generalizations. One of such generalizations is k-submodular functions. In the second part of the talk I will present an application of such functions in computer vision. Based on joint work with Igor Gridchyn, Andrei Krokhin, Michal Rolínek, Johan Thapper and Stanislav Zivný [1, 2, 3].

References

[1] Vladimir Kolmogorov, Johan Thapper and Stanislav Zivný. (2015). The power of linear programming for general-valued CSPs. SIAM Journal on Computing (SICOMP), 44(1):1-36.

[2] Vladimir Kolmogorov, Andrei Krokhin and Michal Rolínek. (2017). The Complexity of General-Valued CSPs. SIAM Journal on Computing (SICOMP), 46(3), 1087-1110.

[3] Vladimir Kolmogorov and Igor Gridchyn. (2015). Potts model, parametric maxflow and k-submodular functions. International Conference on Computer Vision (ICCV).

Convergence rates for an inertial algorithm of gradient type associated to a smooth nonconvex minimization

László Szilárd Csaba

Technical University of Cluj-Napoca
laszloszis zi@yahoo.com

We investigate an inertial algorithm of gradient type in connection with the minimization of a nonconvex differentiable function. The algorithm is formulated in the spirit of Nesterov's accelerated convex gradient method. We show that the generated sequences converge to a critical point of the objective function, if a regularization of the objective function satisfies the Kurdyka-Łojasiewicz property. Further, we provide convergence rates for the generated sequences and the function values formulated in terms of the Łojasiewicz exponent.

Bilevel composite minimization problems

Yura Malitsky

Georg-August-University of Goettingen
y.malitskyi@math.uni-goettingen.de

In this talk we consider a novel diagonal penalty algorithm for solving bilevel composite minimization problem

$$\min_x g_1(x) + f_1(x) \quad s.t. \quad x \in arg \min_y g_2(y) + f_2(y),$$

where $g1, g2$ are convex nonsmooth and $f1, f2$ are convex smooth functions with Lipschitz gradients. As a particular important problem which the above model captures is a nonsmooth minimization over difficult constraints. We survey existing approaches to solve the given problem and discuss their limitations. For the simplest case when the inner problem is just a linear system, we show that our method is equivalent to the primal-dual hybrid gradient method, analyzed by Chambolle and Pock.

Quasi-Newton methods in structured nonsmooth optimization – a hybrid approach

Florian Mannel[1] and Armin Rund[2]

[1−2] *University of Graz, Austria*
florian.mannel@uni-graz.at[1], armin.rund@uni-graz.at[2]

We present an algorithm for the efficient solution of structured nonsmooth operator equations in Banach spaces. Here, the term structured indicates that we consider equations which are composed of a smooth and a semismooth mapping. Equations of this type occur, for instance, as optimality conditions of structured nonsmooth optimization problems such as LASSO in machine learning. The new algorithm combines a semismooth Newton method with a quasi-Newton method. This hybrid approach retains the local superlinear convergence of both these methods under standard assumptions. Since it is known that quasi-Newton methods, in general, cannot achieve superlinear convergence in semismooth settings, this is rather satisfying from a theoretical point of view. The most striking feature of the new method, however, is its numerical performance. On nonsmooth PDE-constrained optimal control problems it is at least an order of magnitude faster than semismooth Newton methods, and these speedups persist when globalization techniques are added. Most notably, the hybrid approach can be embedded in a matrix-free limited-memory truncated trust-region framework to efficiently solve nonconvex and nonsmooth large-scale real-world optimization problems, as we will demonstrate by means of an example from magnetic resonance imaging. In this challenging environment it dramatically outperforms semismooth Newton methods, sometimes by a factor of fifty and more. All of these topics are addressed in the talk.

A proximal minimization algorithm for structured nonconvex and nonsmooth problems

Radu Ioan Bot[1], Ernö Robert Csetnek[2] and Dang-Khoa Nguyen[3]

[1−3] *Faculty of Mathematics, University of Vienna*
radu.bot@univie.ac.at[1], ernoe.robert.csetnek@univie.ac.at[2], dang-khoa.nguyen@univie.ac.at[3]

We propose a proximal algorithm for minimizing objective functions consisting of three summands: the composition of a nonsmooth function with a linear operator, another nonsmooth function, each of the nonsmooth summands depending on an independent block variable, and a smooth function which couples the two block variables. This is an extension of the model considered in [1]. The algorithm is a full splitting method, which means that the nonsmooth functions are processed via their proximal operators, the smooth function via gradient steps, and the linear operator via matrix times vector multiplication. We provide suffiucient conditions for the boundedness of the generated sequence and prove that any cluster point of the latter is a KKT point of the minimization problem. In the setting of the Kurdyka- Lojasiewicz property we show global convergence, and derive convergence rates for the iterates in terms of the Lojasiewicz exponent.

References

[1] J. Bolte, S. Sabach, M. Teboulle. (2014). Proximal alternating linearized minimization for nonconvex and nonsmooth problems. Mathematical Program- ming, 146(1): 459-494.

[2] R. I. Bot, E. R. Csetnek, D.-K. Nguyen. (2018). A proximal mini- mization algorithm for structured nonconvex and nonsmooth problems. arXiv:1805.11056.

Characterisation of feasible bookings in tree-like steady-state gas networks

Martine Labbé[1], Fränk Plein[2] and Martin Schmidt[3]

[1−2] *Département d'Informatique, Université Libre de Bruxelles,*
frank.plein@ulb.ac.be[2], martine.labbe@ulb.ac.be[1]
[3] *Department Mathematik, Friedrich-Alexander-Universität Erlangen-Nürnberg*
mar.schmidt@fau.de

Natural gas is up to the present day one of the main energy sources in Europe. It is therefore crucial to have a good understanding of problems of gas transportation and the functioning of the markets. After recent European directives, transport and trade have been separated in gas markets. A supplier passes contracts with the transportation system operator (TSO), under the form of bookings for maximum injection and withdrawal rights at entry and exit nodes, respectively. On a day-ahead basis, suppliers nominate the quantity of gas to be injected and withdrawn. A nomination within booked capacity rights for which entries and exits are balanced is called booking-compliant. The TSO has to be able to transport all booking-compliant nominations. We recall the definition of feasible nominations in steady-state gas networks, found in the literature. We then introduce the problem of checking the feasibility of bookings, modelled as a robust feasibility problem in which the uncertainty set is the polytope of booking-compliant nominations. Based on a characterisation of feasible nominations in [1], we characterise feasible bookings on tree-like networks composed only of pipelines (i.e. we do not consider controllable elements like compressors, etc.) in terms of inequalities on the optimal values of a family of nonlinear nonconvex optimisation problems over the polytope of booking-compliant nominations. We show that these problems are equivalent to a series of continuous knapsack problems and solve them in polynomial time using dynamic programming.

References

[1] Gotzes, C., Heitsch, H., Henrion, R., Schulz, R. (2016). On the quantification of nomination feasibility in stationary gas networks with random load. Mathematical Methods of Operations Research, 84(2), 427-457.

A variable metric approach for the inexact inertial forward–backward method

S. Rebegoldi[1], S. Bonettini[2] and V. Ruggiero[3]

[1,3] *Dipartimento di Matematica e Informatica, Università degli studi di Ferrara, Via Saragat 1, 44121 Ferrara, Italy*
simone.rebegoldi@unife.it[1], **valeria.ruggiero@unife.it**[3]
[2] *Dipartimento di Scienze Fisiche, Informatiche e Matematiche, Università degli studi di Modena e Reggio Emilia, Via Campi 213/b, 41125 Modena, Italy*
silvia.bonettini@unimore.it[2]

We propose a generalized inertial forward–backward algorithm suited to address structured convex problems, where the objective function is given by the sum of a convex differentiable term and a convex nondifferentiable one [3]. Such algorithm can be considered as a scaled inexact version of the FISTA algorithm [2], in which the proximal operator may be computed inexactly, in order to deal with non-proximable terms, and with respect to a variable metric, with the aim of recovering the acceleration exhibited by FISTA. The projection of the inertial step onto the feasible set is also considered. We prove an $o(1/k^2)$ convergence rate for the function values and the convergence of the iterates generated by the algorithm, under some verifiable assumptions on both the accuracy of the proximal–gradient computation and the variable metric selection [1,3]. The effectiveness of the proposed approach is then validated with a numerical experience on synthetic Poisson data.

References

[1] H. Attouch, and J. Peypouquet, (2016), The rate of convergence of Nesterov's accelerated forward–backward method is actually faster than $1/\mathrm{k}^2$, SIAM Journal on Optimization 26(3), 1824–1834.

[2] A. Beck, and M. Teboulle, (2009), A fast iterative shrinkage–thresholding algorithm for linear inverse problems, SIAM Journal on Imaging Sciences 2(1), 183–202.

[3] S. Bonettini, S. Rebegoldi, and V. Ruggiero, (2018), Inertial variable metric techniques for the inexact forward–backward algorithm. Submitted.

Dynamic Multivariate Programming

Birgit Rudloff[1], Gabriela Kovacova[2] and Zachary Feinstein[3]

[1−2] *Vienna University of Economics and Business*
brudlofft@wu.ac.at[1]
[3] *Washington University in St. Louis*

In this talk, I present a set-valued Bellman principle. As a first example, a famous vector optimization problem is considered: the dynamic mean-risk portfolio optimization problem. Usually, this problem is scalarized and it is well known that this problem does not satisfy the (scalar) Bellman principle. However, when we leave it in its original form as a vector optimization problem, the upper images, whose boundary is the efficient frontier, recurse backwards in time [4]. Conditions are presented under which this recursion can be exploited directly to compute a solution in the spirit of dynamic programming. As a second example, the set-valued Bellman principle appears when considering time consistency of multivariate risk measures. Similar to the first example, one can show that a (fixed) scalarization of the problem does not satisfy the (scalar) Bellman principle, but a particular moving scalarization leads also to scalar time consistency [2, 1]. This is an observation also made in [3], but with our approach this moving scalarization is part of the solution and not needed as an input.

References

[1] Z. Feinstein, B. Rudloff (2018): Time consistency for scalar multivariate risk measures. Submitted for publication.

[2] Z. Feinstein, B. Rudloff (2018): Scalar multivariate risk measures with a single eligible asset. Submitted for publication.

[3] Ch. Karnam, J. Ma, J. Zhang (2017): Dynamic approaches for some timeinconsistent optimization problems. The Annals of Applied Probability, 27(6), 3435 - 3477.

[4] G. Kovacova, B. Rudloff (2018): Time consistency of the mean-risk problem. Submitted for publication.

On the convergence of projection free Hessian Barrier Gradient Algorithms

Mathias Staudigl[1], Panayotis Mertikopoulos[2], Immanuel Bomze[3] and Werner Schachinger[4]

[1] *Maastricht University, Department of Quantitative Economics*
m.staudigl@maastrichtuniversity.nl[1]
[2] *CNRS, Laboratoire d'Informatique de Grenoble*
panayotis.mertikopoulos@imag.fr[2]
[3−4] *University of Vienna, Department of Statistics and Operations Research,*
immanuel.bomze@univie.ac.at[3]**, werner.schachinger@univie.ac.at**[4]

We are interested in solving problems of the form

$$\min f(x) \quad s.t.: \; Ax = b, x \in \bar{C}$$

$\bar{C}$ is the closure of an open convex set $C \subset \mathbb{R}^n$ and A is an $m \times n$ matrix of rank m. We are given a matrix-valued function H which is adapted to the geometry of C. We use this function to define a variable metric given by $\langle u, v \rangle_{H(x)} := u^T H(x) v$ The algorithm we are considering is the recursive scheme

$$x^{k+1} = x^k + \alpha^k P_{x^k} H(x^k)^{-1} \nabla f(x^k)$$

where P_x is a projection. The method above can be seen as a Euler discretization of Riemannian gradient flows on a manifold [1, 2, 3], and its recent followup work [4]. We investigate viability, stability, convergence and complexity of the algorithm, and potentials for acceleration and compare its computational advantage to state-of-the-art first-order methods. In particular, we describe new first-order techniques based on recent development of generalized self-concordant functions [6]. We will also report on recent advances on (HBA) when applied to non-convex, non-smooth problems, with applications to non-convex statistical estimation problems and neural networks. Part of this talk is based on joint recent work together with Immanuel M. Bomze, Panayotis Mertikopoulos and Werner Schachinger [5].

References

[1] Hessian Riemmannian Gradient Flows in Convex Programming; Felipe Alvarez, Jerome Bolte, Olivier Brahic, SIAM J. CONTROL OPTIM, Vol. 43, No. 2, pp. 477-501

[2] Regularized Newton method for unconstrained convex optimization, Roman Polyak, Math. Program., Ser. B (2009) 120, p. 125-145

[3] Barrier Operators and Associated Gradient-Like Dynamical Systems for Constrained Minimization Problems, Marc Teboulle, Jerome Bolte, SIAM J. CONTROL OPTIM, Vol. 42, No. 4, pp. 1266-1292.

[4] A variational perspective on accelerated methods in optimization,Andre Wibisono, Ashia C. Wilson, and Michael I. Jordan, PNAS, 2016

[5] Hessian barrier algorithms for linearly constrained optimization problems, Immanuel M. Bomze, Panayotis Mertikopoulos, Werner Schachinger, Mathias Staudigl, Arxived at https://arxiv.org/abs/1809.09449

[6] Generalized self-concordant functions: a recipe for Newton-type methods. Mathematical Programming, forthcoming, https://doi.org/10.1007/s10107- 018-1282-4

A p-Median Based Exact Method for the Large-Scale Optimal Diversity Management Problem

A. Masone[1], A. Sforza[2], C. Sterle[3], A. Ushakov[4] and I. Vasilyev[5]

[1−3] *Department of Electrical Engineering and Information Technology, University of Naples, 80125 Naples, Italy*

adriano.masone@unina.it[1], sforza@unina.it[2], claudio.sterle@unina.it[3]

[4−5] *Matrosov Institute for System Dynamics and Control Theory, Siberian Branch of the Russian Academy of Sciences, Lermontov str. 134, Irkutsk 664033, Russia*

aushakov@icc.ru[4], vil@icc.ru[5]

The p-median problem (PMP) is the well known network optimization problem of discrete location theory. In many real applications PMP is defined on large scale networks, for which ad-hoc methods have to be developed. An interesting industrial application is constituted by the optimal diversity management problem, ODMP, arising when a company producing a highly customizable good needs to satisfy many client demands with various subset of options, but only a limited number of option configurations can be produced [1]. Exploiting a suitable network representation, ODMP can be formulated as a PMP on a large-scale disconnected network. In this paper we propose an improved decomposition approach where several smaller PMPs related to the network components can be solved instead of the initial large-scale problem [2]. Proposed approach drastically reduces number and dimension of these subproblems, solving them to optimality, and combining their solutions to find the optimal solution of the original problem, formulated as a multiple choice knapsack problem. The computational tests show that it is able to optimally solve known and new test instances (coming from FCA group) up to 110805 nodes and 8153023 arcs, considerably outperforming state-of-the-art approaches.

References

[1] Briant, O., Naddef, D. (2004). The Optimal Diversity Management Problem. Operations Research, 52(4), 515.-526.

[2] Avella, P., Boccia, M., Martino, C.D., Oliviero, G., Sforza, A., Vasilyev, I. (2005). A decomposition approach for a very large scale optimal diversity management problem. 4OR, 3(1), 23.-37.

A dual reoptimization scheme for solving large-scale stochastic optimization problems

Danial Mohseni-Taheri[1], Selva Nadarajah[2] and Alessio Trivella[3]

[1−2] *College of Business Administration, University of Illinois at Chicago*
smohse3@uic.edu[1], selvan@uic.edu[2]
[3] *Department of management engineering, Technical University of Denmark*
atri@dtu.dk[3]

Markov decision processes (MDPs) arising in many applications are high dimensional and subject to well-known curses of dimensionality. A variety of approximate dynamic programming methods have been developed to solve different classes of MDPs. Nevertheless, methods to tackle MDPs which are high dimensional in both the endogenous and exogenous components of their state are limited and rely on problem structure (e.g. convexity) and sophisticated value function approximations. In this paper, we propose a more general framework to approach this intractable problem class. We consider the information relaxation method that is typically used to obtain dual bounds on the MDP optimal value [1], and use it to develop a novel dual reoptimization scheme (DRH) that extracts non-anticipative decision rules from sample action distributions. Specifically, we obtain a distribution of actions at a given MDP stage and state by combining Monte Carlo simulation with the solution of dual mathematical programs on sample paths. We develop some theoretical support for our DRH method, and apply it to an emerging energy real option problem where a company has to construct a dynamic power procurement portfolio to meet a renewable power target by a future date at minimum cost. In other words, the company has to supply a specific percentage of its electricity demand from renewable generators, and has access to multiple short and long-term procurement options. We find that our DRH method outperforms commonly used primal re-optimization methods and simple heuristics on realistic instances. Thus, DRH emerges as a promising framework to tackle high-dimensional MDPs.

References

[1] D. B. Brown, J. E. Smith, and P. Sun. (2010) Information relaxations and duality in stochastic dynamic programs. Operations Research, 58(4):785-801.

Denoising and Completion of Structured Low-Rank Matrices via Iteratively Reweighted Least Squares

Christian Kümmerle[1] and Claudio Mayrink Verdun[2]

[1−2] *Technische Universität München*
christian.kuemmerle@ma.tum.de[1], **verdun@ma.tum.de**[2]

In this work we consider two related problems: first, let $X \in \mathbb{C}^{d_1 \times d_2}$ be a Hankel matrix with $X = (X_{ij}) = x_{i+j-1}$ with some $x \in \mathbb{C}^n, n = d_1 + d_2 - 1$, and consider the task to approximate X by a low-rank Hankel matrix $\hat{Z}$ such that

$$\hat{Z} = arg \min_{\substack{\mathbb{C}^{d_1 \times d_2} \ni Z \, is \, Hankel \\ rank(Z) \leq r}} ||Z - X||^2_{F(w)}$$

where $|| \cdot ||_{F(w)}$ is a suitable weighted Frobenius norm and $r \in \mathbb{N} < \min(d_1, d_2)$. Secondly, let $X = (X_{ij}) = (x_{i+j-1})$ be a Hankel matrix, $\Phi : \mathbb{C}^n \to \mathbb{C}^m$ a subsampling operator, $\mathcal{H} : \mathbb{C}^n \to \mathbb{C}^{d_1 \times d_2}, x \to \mathcal{H}(x) = (\mathcal{H}(x)_{i,j}) = (x_{i+j-1})$ be the linear Hankel operator and let

$$Find \, \hat{Z} = \mathcal{H}(\hat{z}) \, s.t. \, \hat{z} = arg \min_{z \in \mathbb{C}^n, \Phi(z) = \Phi(x)} rank(\mathcal{H}(z)),$$

They arise in fields like parallel MRI, system identification, direction of arrival and interpolation of seismic data. We propose a new Iteratively Reweighted Least Squares (IRLS) algorithm for the problem of completing or denoising low-rank Hankel matrices. The algorithm optimizes an objective based on a non-convex surrogate of the rank by solving a sequence of quadratic problems. Our strategy combines computational efficiency, as it operates on a lower dimensional generator space of the structured matrices, with high statistical accuracy which can be observed in experiments on hard estimation and completion tasks. Our experiments show that the proposed Structured HM-IRLS exhibits an empirical recovery probability close to 1 from fewer samples than the state-of-the-art in a Hankel matrix completion task arising from the problem of spectral super-resolution of badly separated frequencies.

References

[1] Kümmerle, C. and Sigl, J.. Harmonic Mean Iteratively Reweighted Least Squares for Low-Rank Matrix Recovery. To appear in Journal of Machine Learning Research

[2] Kümmerle, C. and Verdun, C. M.. Non-Convex Spectral Compressed Sensing of Non-Separated Signals. Preprint in preparation, 2018.

Variational mesh denoising and surface fairing using the total variation of the normal

J. Vidal-Nunez[1], R. Herzog[2], R. Bergmann[3], S. Schmidt[4] and M. Herrmann[5]

[1−3] *Faculty of Mathematics, TU Chemnitz*
jose.vidal-nunez@mathematik.tu-chemnitz.de,[1],
roland.herzog@mathematik.tu-chemnitz.de[2],
ronny.bergmann@mathematik.tu-chemnitz.de[3]
[4−5] *2Faculty of Mathematics, University of Wuerzburg*
stephan.schmidt@mathematik.uni-wuerzburg.de[3],
marc.herrmann@mathematik.uni-wuerzburg.de[4]

In this talk we present a novel approach to solve a surface mesh denoising problem. The goal is to remove noise while preserving shape features such as sharp edges. The total variation of the normal vector field in various norms serves as a regularizing functional to achieve this goal, see also [1] and [2]. We propose a novel technique for the numerical treatment of this problem, based on a level-set representation of the surface and Galerkin finite element discretization.

References

[1] Wu, X., Zheng, J., Cai, Y., and Fu C-F. (2015) Mesh denoising using extended ROF model with l1 fidelity. Computer Graphics Forum, 34(7), 35–45.

[2] Zhang, H., Wu, C., Zhang, J., and Deng, J. (2015) Variational mesh denoising using total variation and piecewise constant function space. IEEE Transactions on Visualization and Computer Graphics, 21(7), 873–886.

Optimization algorithms based on operator splitting and energy-preserving numerical integrators

Adrian Viorel

Department of Mathematics, Technical University of Cluj-Napoca
Aadrian.Viorel@math.utcluj.ro

The present contribution deals with nonconvex optimization from a dynamical systems perspective. More precisely, we consider the second order evolution equation

$$\ddot{u} + \dot{u} = -\nabla G(u) \tag{1}$$

which can be understood (cf. [1]) as describing a damped mechanical system with the nonconvex objective function G as the system's potential energy. Its asymptotic behavior has been analyzed quite recently by Bégout et. al. who show the convergence of solutions, at precise rates, to minimizers of G. Our aim is to investigate a discrete dynamical system which faithfully reproduces the properties of (1). The discrete dynamical system is obtained form (1) by means of a two-stage discretization consisting of a Strang splitting semidiscretization [3, 2] that separates the conservative and dissipative parts of the system followed by an energy-preserving numerical integration [4] of the conservative part. Furthermore, the same approach can be employed for the variable damping model

$$\ddot{u} + \frac{\gamma}{t}\dot{u} = -\nabla G(u),$$

which can be regarded as a continuous version of the Nesterov algorithm (see [5]).

References

[1] H. Attouch, X. Goudou, and P. Redont. The heavy ball with friction method. I. The continuous dynamical system: global exploration of the local minima of a real-valued function by asymptotic analysis of a dissipative dynamical system. Commun. Contemp. Math., 2(1):1–34, 2000.

[2] R. Glowinski and S. J. Osher, editors. Splitting methods in communication, imaging, science, and engineering. Scientific Computation. Springer, Cham, 2016.

[3] E. Hansen, F. Kramer, and A. Ostermann. A second-order positivity preserving scheme for semilinear parabolic problems. Appl. Numer. Math., 62(10):1428–1435, 2012.

[4] G. R. W. Quispel and D. I. McLaren. A new class of energy-preserving numerical integration methods. J. Phys. A, 41(4):045206, 7, 2008.

[5] W. Su, S. Boyd, and E. J. Candès. A differential equation for modeling Nesterov's accelerated gradient method: theory and insights. J. Mach. Learn. Res., 17:Paper No. 153, 43, 2016.

ROSSAUER LÄNDE
THE FACULTY
(WORKSHOP VENUE)
Hahngasse
Roßauer Lände
Währinger Str.
Berggasse
Türkenstr.
Hörlgasse
Untere Augartenstr
Roßauer Kaserne
Maria-Theresien-Str.
Schottenring
SCHOTTENRING
Universitätsstr.
SCHOTTENTOR
Main Building
Universitätsring
Franz-Josefs-Ka

The Vienna Graduate School on Computational Optimization is a joint initiative between the University of Vienna and the TU Wien funded by the Austrian Science Funds (FWF).