Optimization, Game Theory, and Data Analysis: Workshop Program

Schedule

Date	Time	Session
20.12.	13:00 - 13:15	Welcome
	13:15 - 15:15	**Copositive Optimization**
	15:15 - 15:45	Coffee Break
	15:45 - 16:45	**Evolution and Games**
	17:00 - 17:30	Addresses and Music
	17:30 - 19:30	Buffet Dinner
	19:30 - 21:00	City Walk
21.12.	09:00 - 10:30	**Data Science**
	10:30 - 11:00	Coffee Break
	11:00 - 12:30	**Quadratic Optimization**
	12:30 - 13:45	Lunch Break
	13:45 - 14:45	**Uncertainty and Multi-Criteria Optimization**
	14:45 - 16:15	Panel Discussion: "Short-lived hypes and failures in OR"
	16:15 - ...	Closing, Coffee

Session: Copositive Optimization

- Paula Amaral (FCT and CMA, Universidade Nova de Lisboa, Portugal)
  Copositivity in fractional optimization
- Peter Dickinson (Rabobank, The Netherlands)
  Generating irreducible copositive matrices using the stable set problem
- Mirjam Dür (Universität Augsburg, Germany)
  Copositive optimization and completely positive matrix factorization
- Naomi Shaked-Monderer (Max Stern Yezreel Valley College, Israel)
  Some copositive thoughts

Session: Evolution and Games

- Anita Schöbel (Universität Göttingen, Germany): The search for the tree of life: A location problem in the phylogenetic tree space
- Christina Pawlowitsch (Université Panthéon-Assas, Paris 2, France)
  Costly signals: rationality and evolution (a classification)

Session: Data Science

- Samuel Rota Bulò (Mapillary Research, Austria): Infection and Immunization Dynamics and its application to data clustering
- Annabella Astorino (ICAR-CNR, Italy)
  DC models for machine learning problems
- Paula Brito (Universidade do Porto, Portugal)
  Linear models for complex data analysis

Session: Quadratic Optimization

- Laura Palagi (Sapienza Università di Roma, Italy)
  Fast computation of bounds in quadratic integer programming
- Veronica Piccialli (Università degli Studi di Roma "Tor Vergata", Italy)
  A new branch-and-bound algorithm for standard quadratic programming problems
- E. Alper Yıldırım (Koç Üniversitesi, Turkey)
  Global solutions of nonconvex standard quadratic programs via mixed integer linear programming formulations

Session: Uncertainty and Multi-Criteria Optimization

- Gabriele Eichfelder (TU Ilmenau, Germany)
  Using convex underestimators for global multiobjective optimization
- Francesca Maggioni (Università degli Studi di Bergamo, Italy)
  Sampling methods for multistage robust convex optimization problems

Panel Discussion: "Short-lived hypes and failures in OR"

- Florian Jarre (Heinrich Heine Universität Düsseldorf)
- Joaquim Júdice (Universidade de Coimbra)
- Abdel Lisser (Université Paris Sud)
- Ivana Ljubic (ESSEC Paris)
- Marco Locatelli (Università degli Studi Parma)
- Fabio Schoen (Università degli Studi Firenze)
- Fabio Tardella (Sapienza Università di Roma)
- Tamás Terlaky (Lehigh University, Bethlehem PA)

Abstracts (speakers in alphabetical order)

Copositivity in fractional optimization (Paula Amaral)

Abstract: In this talk I will present some of the work developed in collaboration with Immanuel Bomze. We focused mainly in fractional problems involving quadratic functions and developed conic formulations for these hard optimization problems. The practical contributions are related with relaxations of these conic formulations that yield efficient lower bounds. From fractional quadratic functions in continuous and binary variables to continuous minimax fractional problems, the results were always encouraging and contribute to certify the well-known quality of bounds obtained from relaxations of copositive reformulations. This talk is also the testimony of a friendship that developed over these years.

DC models for machine learning problems (Annabella Astorino)

Abstract: In the last twenty years, several scientists coming from different subfields of numerical optimization have entered the area of data analysis and machine learning, mainly tackling problems of clustering and classification type. Machine learning is, in fact, a data-driven inference where the objective is to predict known properties (supervised learning) or to discover unknown characteristics (unsupervised learning).

This work deals with pattern classification, which consists in categorizing data into different classes on the basis of their similarities. From such a general definition, several mathematical formulations have stemmed, with application in many fields. Some well known examples are object recognition in machine vision, gene expression profile analysis, DNA and protein analysis and many others.

From the mathematical point of view, classification reduces to finding separation surfaces in the sample space, where the objects (samples) are represented through their attributes. Consequently, convexity and separation properties of sets (e.g. theorems such as Hahn-Banach's) play a relevant role. In practical cases, since it is not always easy to check whether the theoretical conditions of existence of such surfaces hold, it is necessary to tackle the classification-separation problems by means of tools which are able to provide a quantitative answer even when the sets are not separable. In this case, it is possible to adopt a numerical optimization model to minimize some misclassification measure.

The area, in addition to traditional nonlinear programming techniques (such as quadratic programming), can benefit from many methodologies of nondifferentiable optimization, where problems of the minmax and minmaxmin type are tackled. More specifically, the nonconvex and nonsmooth nature of the misclassification objective function can be handled by means of nonsmooth DC optimization techniques, where functions are represented as a difference of two convex (DC) functions. Nonsmooth DC optimization, in fact, constitutes an important and broad subclass of nonconvex optimization, particularly suitable and useful for many real-world applications, well beyond Machine Learning.

In this presentation we introduce some classification problems of supervised and semi-supervised type by proposing for them possible DC formulations. In particular, we present the non linear separation problems of spherical, polyhedral and min-max type, the Transductive Support Vector Machine (TSVM) approach for semisupervised classification and, finally, the binary Multiple Istance Learning (MIL) classification problem.

Linear models for complex data analysis (Paula Brito)

Abstract: In classical Statistics and Multivariate Data Analysis data is usually represented in a matrix where each row represents a statistical unit, or "individual", for which one single value is recorded for each numerical or categorical variable (in columns). This representation model is however too restricted when the data to be analysed comprises variability. That is the case when the entities under analysis are not single elements, but groups formed on the basis of some given common properties and the observed variability within each group should be taken into account. To this aim, new variable types have been introduced, whose realizations are not single real values or categories, but sets, intervals, or distributions over a given domain. Symbolic Data Analysis provides a framework for the representation and analysis of such data, taking into account their inherent variability. In this talk, we consider the case of aggregate numerical data described by empirical distributions, known as histogram data. Linear models for such distributional variables are proposed, which rely on the representation of histograms by the associated quantile functions. These then allow for linear regression as well as for linear discriminant analysis for histogram-valued data. An application of the proposed methodology will be presented.

This is joint work with Sónia Dias and Paula Amaral.

Generating irreducible copositive matrices using the stable set problem (Peter Dickinson)

Abstract: A matrix is defined to be copositive if its quadratic product with all nonnegative vectors is itself nonnegative. In this talk, it is considered how simple graphs can be used to generate copositive matrices, and necessary and sufficient conditions are given on graphs for the generated matrices to then have the property of being irreducible with respect to the set of positive semidefinite plus nonnegative matrices. By considering these conditions for graphs with at most thirteen vertices, tens of thousands of irreducible copositive matrices are generated, and remarkably many of them turn out to be extreme copositive matrices.

This talk is based on joint work with Reinier de Zeeuw.

Copositive optimization and completely positive matrix factorization (Mirjam Dür)

Abstract: A copositive optimization problem is a problem in matrix variables with a constraint which requires that the matrix be in the copositive cone. This means that its quadratic form takes nonnegative values over the nonnegative orthant. Many combinatorial problems like for example the maximum clique problem can be equivalently formulated as a copositive problem. Burer (2009) showed that also any nonconvex quadratic problem with linear constraints and binary variables can be reformulated as such a copositive problem. This is remarkable, since by this approach, a nonconvex problem is reformulated equivalently as a convex problem. The complexity of the original problem is entirely shifted into the cone constraint. The dual of a copositive problem involves the cone of completely positive matrices, that is, matrices that can be factorized as $A = BB^T$ with an entrywise nonnegative matrix B. Given the matrix A, finding such a factorization is a nontrivial task but necessary in order to recover the solution of the underlying combinatorial problem. In the talk, we will discuss recent advances in this area.

Using convex underestimators for global multiobjective optimization (Gabriele Eichfelder)

Abstract: Standard techniques in global optimization are d.c. decompositions and branch-and-bound procedures based on subdivision strategies of the search domain. Lower bounds on subsets of the feasible set combined with upper bounds derived from objective function values are used to obtain discarding tests. Based on these standard techniques a well-known and efficient method in single-valued global optimization is the alpha Branch and Bound (alphaBB) method which uses convex underestimators of the objective function.

In multiobjective optimization one studies optimization problems with multiple objectives. Applying the above techniques results in convex multiobjective optimization sub-problems which deliver lower bound sets instead of a single scalar lower bound. Also the upper bounds are sets of points now. Hence, sets have to be compared. To make this numerically tractable we use outer approximations from convex multiobjective optimization combined with the concept of local upper bounds, which was introduced for combinatorial multiobjective optimization problems. We recall the mentioned techniques and propose based on that a numerical algorithm for solving nonconvex multiobjective optimization problems globally.

This talk is based on joint work with Julia Niebling.

Sampling methods for multistage robust convex optimization problems (Francesca Maggioni)

Abstract: In this talk, probabilistic guarantees for constraint sampling of multistage robust convex optimization problems are derived. The dynamic nature of these problems is tackled via the so-called scenario-with-certificates approach allowing to avoid the conservative use of explicit parametrizations through decision rules. An explicit bound on the probability of violation of the randomized solution is provided. Numerical results dealing with a multistage inventory management problem show the efficacy of the proposed approach.

This is joint work with Fabrizio Dabbene and Georg Pflug.

Fast computation of bounds in quadratic integer programming (Laura Palagi)

Abstract: In this talk we present branch-and-bound algorithms that generalize the approach for unconstrained nonconvex quadratic integer programming proposed by Buchheim, De Santis, Palagi, Piacentini [SIOPT, (2013), 23(3), pp. 1867-1889] to the presence of linear constraints. The main feature of the latter approach consists of a sophisticated preprocessing phase, leading to a fast enumeration on the branch-and-bound nodes. The focus of the talk will be on how lower bounds can be computed efficiently exploiting the structure of the problem. Experimental results on randomly generated instances show that the approach significantly outperforms the MIQP solver of CPLEX 12.6 for instances with a small number of constraints. Insights on how to extend the approach for solving the nonconvex quadratic knapsack problem will be presented.

Costly signals: rationality and evolution (a classification) (Christina Pawlowitsch)

Abstract: My talk will be based on current work on a book with Josef Hofbauer. In this book we investigate equilibrium refinement from both a classical and an evolutionary point of view in three simple classes of discrete costly-signaling games. Discrete models are empirically relevant and they are an important theoretical reference point for they allow one to discern in a clear-cut way under which conditions which kind of signaling equilibria can exist. I will focus on three questions:

- Is costly signaling necessarily a waste of resources or can it improve social well-being?
- Is costly signal necessarily "honest" - or can there be "cheating"?
- What is meaning-making in a costly signaling game?

I will illustrate our results in a model of strategic language use.

This is joint work with Josef Hofbauer.

A new branch-and-bound algorithm for standard quadratic programming problems (Veronica Piccialli)

Abstract: In this work we propose convex and LP bounds for standard quadratic programming (StQP) problems and employ them within a branch-and-bound

approach. We first compare different bounding strategies for StQPs in terms both of the quality of the bound and of the computation times. It turns out that the polyhedral bounding strategy is the best one to be used within a branch-and-bound scheme. Indeed, it guarantees a good quality of the bound at the expense of a very limited computation time. The proposed branch-and-bound algorithm performs an implicit enumeration of all the KKT (stationary) points of the problem. We compare different branching strategies exploiting the structure of the problem. Numerical results on randomly generated problems (with varying density of the underlying convexity graph) are reported which show the effectiveness of the proposed approach, in particular in limiting the growth of the number of nodes in the branch-and-bound tree as the density of the underlying graph increases. Finally, we introduce a new branch and bound method for StQP arising in a game theory setting with a specific structure and report some preliminary results.

This is joint work with Giampaolo Liuzzi and Marco Locatelli.

The search for the tree of life: a location problem in the phylogenetic tree space (Anita Schöbel)

Abstract: We describe a problem arising in computational biology which so far has been researched mainly in statistics: Locating a tree in the phylogenetic tree space. We show that this problem can be modeled as a location problem in a metric space, the so-called tree space. The tree space consists of regions in which distances are measured Euclidean, but traveling between the different regions has similarities with the distances in location problems with barriers.

After introducing the problem we show which types of location problems arise and how results of location theory can be used to identify good phylogenetic trees. We present bounds on the optimal objective value and a blockwise coordinate descent approach which uses the local Euclidean structure of the tree space and which converges to an optimal solution.

This is joint work with Marco Botte.

Some copositive thoughts (Naomi Shaked-Monderer)

Abstract: Some results on copositive matrices with spectral radius and all negative eigenvalues equal to $-\rho$: characterization, existences, and copositivity of their Moore-Penrose inverse. These results were motivated by a known sufficient condition for the copositivity of the inverse of a copositive matrix, and an observation made by Bomze in a 2008 paper, regarding copositive matrices with spectral radius ρ and an eigenvalue equal to $-\rho$.

Infection and Immunization Dynamics and its application to data clustering (Samuel Rota Bulò)

Abstract: The talk introduces a family of game dynamics called Infection and Immunization Dynamics (InfImmDyn) that I have jointly developed with Immanuel Bomze. InfImmDyn builds upon a central paradigm of evolutionary

game theory, namely the invasion barrier. For partnership games, InfImmDyn exhibits a better asymptotic behavior compared to other popular procedures like Fictitious Play and Replicator Dynamics. Moreover, it features support separation in finite time, which can never be achieved by any interior-point method like those mentioned above. The second part of the talk introduces a game-theoretic framework for data clustering and describes how InfImmDyn and its support-separation properties can be exploited to iteratively extract clusters.

This is joint work with Immanuel Bomze and Marcello Pelillo.

Global solutions of nonconvex standard quadratic programs via mixed integer linear programming formulations (E. Alper Yıldırım)

Abstract: A standard quadratic program is an optimization problem that consists of minimizing a (nonconvex) quadratic form over the unit simplex. We focus on reformulating a standard quadratic program as a mixed integer linear programming problem. We propose two alternative mixed integer linear programming formulations. Our first formulation is based on casting a standard quadratic program as a linear program with complementarity constraints. We then employ binary variables to linearize the complementarity constraints. For the second formulation, we first derive an overestimating function of the objective function and establish its tightness at any global minimizer. We then linearize the overestimating function using binary variables and obtain our second formulation. For both formulations, we propose a set of valid inequalities. Our extensive computational results illustrate that the proposed mixed integer linear programming reformulations significantly outperform other global solution approaches. On larger instances, we usually observe improvements of orders of magnitude.

This is joint work with Jacek Gondzio.