Enforcing Convergence to all Members of the Broyden Family of Methods for Unconstrained Optimization

Mehiddin Al-Baali
Sultan Qaboos University, Oman
albaali@squ.edu.om

The Broyden family of quasi-Newton methods for unconstrained optimization will be considered. It is well-known that if a member of this family is defined sufficiently close to that of the robust BFGS method, then useful theoretical and numerical properties are obtained. These properties will be extended to all members of the above family provided that the current points are sufficiently close to the solution of a convex optimization problem and that the Hessian approximations satisfy certain conditions. A possibility for enforcing these properties to all members of the family will be provided for any starting point and any initial positive definite Hessian approximation. Numerical results will be described to illustrate that some robust, inefficient and divergent Broyden family methods are enforced to be competitive with the standard BFGS method.

Second-order optimality conditions and shooting algorithms for optimal control problems

Joseph Frédéric Bonnans

Inria-Saclay and CMAP, Ecole Polytechnique

Frederic.Bonnans@inria.fr

We will discuss the second-order optimality conditions, both necessary and sufficient, for optimal control problems of ODEs with control and state constraints, for weak and strong solutions. In the second case the second-order optimality conditions typically involve Pontryagin multipliers. We will show the close link with the well-posedness of the shooting algorithm in various settings, and with the analysis of discretization errors.

Libraries of algorithms for solving large-scale LP problem and for global optimization

Yury Evtushenko

Dorodnicyn Computing Centre of RAS, Russia

evt@ccas.ru

LP projection method is used for finding normal solution of LP problem. This approach is close to the quadratic penalty function method and to the modified Langrangian function method. This method yields the exact projection of a given point on the solution set of primal LP problem as a result of the single unconstrained maximization of an auxiliary piecewise quadratic concave function for any sufficiently large values of the penalty parameter. A generalized Newton method with a stepsize chosen using Armijo's rule was used for unconstrained maximization. The proof of globally convergent finitely terminating generalized Newton method for piecewise quadratic function was giving. LP projection method solves LP problems with a very large ($\sim 10^7$) number of variables and moderate ($\sim 10^5$) number of constraints. In a similar way, the exact projection of a given point on the solution set of the dual LP problem can be obtained by nonnegative constrained maximization of auxiliary quadratic function for sufficiently large but finite values of the penalty parameter. A library of algorithms for global optimization was developed on the basis of nonuniform space covering technique which was proposed by author in 1971 for Lipshitzian functions. Later this approach was generalized and was applied to nonlinear programming problem and to multiobjective optimization. The objective and constraints functions may contain both continuous and integer variables and the objective may be non-convex and multiextremal. The covering algorithm for multiobjective optimization has been implemented in the BNB-Solver framework. The BNB-Solver is a generic framework for implementing optimization algorithms on serial and parallel computers. Computational experiments for various test problems demonstrated that this algorithm reliably constructed the ϵ-Pareto set approximations in a reasonable time. Obtained approximations provide uniform covering of the Pareto-set.

Smoothing Approach to Equilibrium Problems with Shared Complementarity Constraints

Masao Fukushima

Kyoto University, Japan

fuku@i.kyoto-u.ac.jp

The equilibrium problem with equilibrium constraints (EPEC) can be looked on as a generalization of Nash equilibrium problem (NEP) and the mathematical program with equilibrium constraints (MPEC). In this paper, we particularly consider a special class of EPECs where a common parametric P-matrix linear complementarity system is contained in all players' strategy sets. After reformulating the EPEC as an equivalent nonsmooth NEP, we use a smoothing method to construct a sequence of smoothed NEPs that approximate the original problem. We consider two solution concepts, global Nash equilibrium and stationary Nash equilibrium, and establish some results about the convergence of approximate Nash equilibria. Moreover we show some illustrative numerical examples.

(Joint work with Ming Hu.)

Cubic overestimation and secant updating for unconstrained optimization

Andreas Griewank

Humboldt Universität, Gremany

griewank@math.hu-berlin.de

The discrepancy between an objective function f and its local quadratic model $f(x) + g(x)^T s + sH(x)s/2 \approx f(x + s)$ at the current iterate x is estimated using a *cubic* term $q\|s\|^3/3$. Potential steps are chosen such that they minimize the overestimating function $g(x)^\top s + s^\top B s/2 + q\|s\|^3/3$ with $B \approx H(x)$. This ensures $f(x+s) < f(x)$ unless B differs significantly from $H(x)$ or the scalar $q > 0$ is too small. Either or both quantities are updated over unsuccessful and successful steps alike. For an algorithm employing both the symmetric rank one update and the BFGS formula we show that either $\inf \|g\| = 0$ or $\sup \|B\| = \infty$, provided the Hessian $H(x)$ is locally Lipschitz on R^n.

MPECs in Function Space

Michael Hintermueller

Humboldt Universität, Germany

hint@math.hu-berlin.de

Several classes of mathematical programs with equilibrium constraints (MPECs), which are posed in function space and which are related to the optimal control of (partial differential) variational inequalities will be considered. Based on set-valued analysis tools suitable stationarity principles will be derived. Particular focus will be put on constraints on the gradient of the state as well as on upper level control constraints. Several extensions of the framework are discussed and a solution algorithm along with numerical results will be presented.

Domain Decomposition techniques in Topology Optimization

Michal Kocvara

School of Mathematics University of Birmingham UK

m.kocvara@bham.ac.uk

Recent advances in the application of domain decomposition techniques to topology optimization of mechanical structures will be presented. The topology optimization problem can be modelled as a (very-)large-scale convex constrained optimization problem. The large dimensionality stems from very fine discretization by finite elements. We will present several ways how to use domain decomposition techniques, known from the solution of linear systems, to solve large instances of the problem. (Joint work with M. Kojima (Tokyo), D. Loghin and J. Turner (Birmingham).)

Subgradient methods for huge-scale optimization problems

Yurii Nesterov

Université Catholique de Louvain, Belgium

Yurii.Nesterov@uclouvain.be

We consider a new class of huge-scale problems, the problems with sparse subgradients. The most important functions of this type are piece-wise linear. For optimization problems with uniform sparsity of corresponding linear operators, we suggest a very efficient implementation of subgradient iterations, which total cost depends logarithmically in the dimension. This technique is based on a recursive update of the results of matrix/vector products and the values of symmetric functions. It works well, for example, for matrices with few nonzero diagonals and for max-type functions.

We show that the updating technique can be efficiently coupled with the simplest subgradient methods, the unconstrained minimization method by B.Polyak, and the constrained minimization scheme by N.Shor. Similar results can be obtained for a new nonsmooth random variant of a coordinate descent scheme. We present also the promising results of preliminary computational experiments.

Detecting Critical Subsets (nodes, edges, shortest paths, or cliques) in Large Networks

Panos M. Pardalos
University of Florida, USA
pardalos@ufl.edu

In network analysis, the problem of detecting subsets of elements important to the connectivity of a network (i.e., critical elements) has become a fundamental task over the last few years. Identifying the nodes, arcs, paths, clusters, cliques, etc., that are responsible for network cohesion can be crucial for studying many fundamental properties of a network. Depending on the context, finding these elements can help to analyze structural characteristics such as, attack tolerance, robustness, and vulnerability. Furthermore we can classify critical elements based on their centrality, prestige, reputation and can determine dominant clusters and partitions. From the point of view of robustness and vulnerability analysis, evaluating how well a network will perform under certain disruptive events plays a vital role in the design and operation of such a network. To detect vulnerability issues, it is of particular importance to analyze how well connected a network will remain after a disruptive event takes place, destroying or impairing a set of its elements. The main goal is to identify the set of critical elements that must be protected or reinforced in order to mitigate the negative impact that the absence of such elements may produce in the network. Applications are typically found in homeland security, energy grid, evacuation planning, immunization strategies, financial networks, biological networks, and transportation. From the member-classification perspective, identifying members with a high reputation and influential power within a social network could be of great importance when designing a marketing strategy. Positioning a product, spreading a rumor, or developing a campaign against drugs and alcohol abuse may have a great impact over society if the strategy is properly targeted among the most influential and recognized members of a community. The recent emergence of social networks such as Facebook, Twitter, LinkedIn, etc. provide countless applications for problems of critical-element detection. In addition, determining dominant cliques or clusters over different industries and markets via critical clique detection may be crucial in the analysis of market share concentrations and debt concentrations, spotting possible collusive actions or even helping to prevent future economic crises. This presentation surveys some of the recent advances for solving these kinds of problems including heuristics, mathematical programing, dynamic programing, approximation algorithms, and simulation approaches. We also summarize some applications that can be found in the literature and present further motivation for the use of these methodologies for network analysis in a broader context.

(Joint work with Steffen Rebennack, Ashwin Arulselvan, Clayton Commander, Vladimir Boginski, Chrysafis Vogiatzis, Jose L. Walteros, Neng Fan, Donatella Granata, and Olga Khvostova)

Complex Optimization in Quantum Physics

Liqun Qi

The Hong Kong Polytechnic University, Hong Kong

maqilq@polyu.edu.hk

In the study of the quantum entanglement problem, there are optimization problems involving complex variables. In this talk, we will analyze such optimization problems and report the result on the minimum Hartree value.

Alternating Linearization for Structured Regularization Problems

Andrzej Ruszczynski
Rutgers University, New Jersey, USA
rusz@business.rutgers.edu

We adapt the alternating linearization method for proximal decomposition to structured regularization problems, in particular, to the generalized lasso problems. The method is related to two well-known operator splitting methods, the Douglas–Rachford and the Peaceman–Rachford method, but it has descent properties with respect to the objective function. Its convergence mechanism is related to that of bundle methods of nonsmooth optimization. We also discuss implementation for very large problems, with the use of specialized algorithms and sparse data structures. Finally, we present numerical results for several synthetic and real-world examples, including a three-dimensional fused lasso problem, which illustrate the scalability, efficacy, and accuracy of the method.

Preconditioning Techniques for Optimization with PDEs

Ekkehard Sachs
University of Trier, Germany
sachs@uni-trier.de

We review recent progress in the numerical solution of optimization problems with partial differential equations as constraints. Preconditioning the Karush-Kuhn-Tucker system in the solution of the necessary optimality system is an important task. We show research efforts in this field during the past years with special emphasis to the role of Bramble-Pasciak preconditioners. We conclude with remarks on the impact of the original nonlinear structure of the problem on preconditioning techniques.

Inexact decomposition methods for constrained convex optimization problems and application to network equilibrium problems

Marco Sciandrone

Università di Firenze, Italy

marco.sciandrone@iasi.cnr.it

In this work we propose convergent inexact decomposition methods for smooth problems defined on the Cartesian product of convex sets. More specifically, we present a conceptual model of decomposition algorithm using a restriction of the feasible sets (in order to possibly exploiting column generation strategies), and gradient projection mappings (for computing inexact solutions of the generated subproblems). The global convergence of the algorithm is proved under suitable assumptions on the restriction of the feasible sets. Due to the generality of the assumptions stated, the proposed algorithm model can be the framework to develop decomposition algorithms for different classes of problems. In the talk we focus on the class of network equilibrium problems, we present convergent decomposition schemes derived from the algorithm model, and we show computational results obtained on medium-large dimensional problems.

The Cubic Regularization Algorithm and Recent Complexity Issues for Nonconvex Optimization

Philippe Toint
University of Namur, Belgium
philippe.toint@fundp.ac.be

The talk will survey recent developments in the analysis of worst-case complexity bounds for algorithms intended for solving nonconvex continuous optimization problems. The convergence to first- and second-order critical points in the unconstrained case will be considered first, and methods such as steepest descent, Newton and several of its variants will be revisited. The talk will also present some approaches for the constrained case. Some relatively surprising results will be given and the special nature of the cubic regularization method (ARC) will be pointed out.

Multilevel Methods for PDE-Constrained Optimization involving Adaptive Discretizations and Reduced Order Models

Stefan Ulbrich

Technische Universität Darmstädt, Germany

ulbrich@mathematik.tu-darmstadt.de

We consider optimization problems governed by time-dependent partial differential equations. Multilevel techniques use a hierarchy of approximations to this infinite dimensional problem and offer the potential to carry out most optimization iterations on comparably coarse discretizations. In this talk we discuss the efficient interplay between the optimization method, adaptive discretizations of the PDE, reduced order models derived from these discretizations, and error estimators. To this end, we describe an adaptive multilevel SQP method that generates a hierarchy of adaptive discretizations during the optimization iteration using adaptive finite-element approximations and reduced order models such as POD. The adaptive refinement strategy is based on a posteriori error estimators for the PDE-constraint, the adjoint equation and the criticality measure. The resulting optimization methods allows to use existing adaptive PDE-solvers and error estimators in a modular way. We demonstrate the efficiency of the approach by numerical examples.

Sparse Hessian Recovery and Trust-Region Methods based on Probabilistic Models

Luis Nunes Vicente

University of Coimbra, Portugal

lnv@mat.uc.pt

In many application problems in optimization, one has little or no correlation between problem variables, and such (sparsity) structure is unknown in advance when optimizing without derivatives. We will show that quadratic interpolation models computed by l1-minimization recover the Hessian sparsity of the function being modeled, when using random sample sets. Given a considerable level of sparsity in the unknown Hessian of the function, such models can achieve the accuracy of second order Taylor ones with a number of sample points (or observations) significantly lower than $O(n^2)$.

The use of such modeling techniques in derivative-free optimization led us to the consideration of trust-region methods where the accuracy of the models is given with some positive probability. We will show that as long as such probability of model accuracy is over 1/2, one can ensure, almost surely, some form of convergence to first and second order stationary points.

The Simplex and Policy-Iteration Methods are Strongly Polynomial for the Markov Decision Problem with a Constant Discount Factor

Yinyu Ye

Stanford University, California, USA

yinyu-ye@stanford.edu

We prove that the classic policy-iteration method (Howard 1960), including the Simplex method (Dantzig 1947) with the most-negative-reduced-cost pivoting rule, is a strongly polynomial-time algorithm for solving the Markov decision problem (MDP) with a constant discount factor. Furthermore, the computational complexity of the policy-iteration method (including the Simplex method) is superior to that of the only known strongly polynomial-time interior-point algorithm for solving this problem. The result is surprising since the Simplex method with the same pivoting rule was shown to be exponential for solving a general linear programming (LP) problem, the Simplex (or simple policy-iteration) method with the smallest-index pivoting rule was shown to be exponential for solving an MDP regardless of discount rates, and the policy-iteration method was recently shown to be exponential for solving a undiscounted MDP. We also extend the result to solving MDPs with sub-stochastic and transient state transition probability matrices.

The Frank-Wolfe Algorithm, Away Steps, and Core Sets in Optimization

E. Alper Yildirim
Koc University, Turkey
alperyildirim@ku.edu.tr

Recently, many studies have centered around developing algorithms for large-scale optimization problems by identifying a small subset of the constraints and/or variables and solving the resulting smaller optimization problem instead. Such small subsets are known as â*core sets.*For a certain class of optimization problems, one can explicitly compute such a small subset with the property that the resulting optimal solution is a close approximation of the optimal solution of the original problem. Such problems mainly include geometric optimization problems such as minimum containment, clustering, and classifcation. In this talk, we discuss the connections between core sets and the Frank-Wolfe algorithm and its variants. In particular, we discuss how the Frank-Wolfe algorithm forms a basis for the existence of small core sets for various large-scale optimization problems.

Hybrid algorithm for power maximization interference alignment problem of MIMO channels

Ya-xiang Yuan

Chinese Academy of Sciences, China

yyx@lsec.cc.ac.cn

Considering a K-user MIMO interference channel, we present a desired signal power maximization model with interference alignment (IA) conditions as constraints to design proper users' precoders and decoders, which forms a complex matrix optimization problem. Courant penalty function technique is applied to combine the leakage interference and the desired signal power as the new objective function. We propose an $e\pm$client hybrid algorithm to solve the problem, which includes two parts. As the first part, a new algorithm iterates with Householder transformation to preserve the orthogonality of precoders and decoders. In each iteration, the matrix optimization problem is decomposed to several one dimensional optimization problems. From any initial point, this algorithm can obtain precoders and decoders with low leakage interference in short time. In the second part, an alternating minimization algorithm with Courant penalty function technique is proposed to perfectly align the interference from the output point of the first part. Analysis shows that generally the hybrid algorithm has lower computational complexity than the existed maximum signal power (MSP) algorithm. Simulations reveal that the hybrid algorithm achieves similar performance as MSP algorithm with less computing time, and shows better performance than the conventional IA algorithm in terms of sum rate.