PREFACE

Series B

Special Issue: Hierarchical Optimization

Kristin P. Bennett[1] · Michael C. Ferris[2] · Jong-Shi Pang[3] · Mikhail V. Solodov[4] · Stephen J. Wright[2]

Hierarchical optimization (HO) is the subfield of mathematical programming in which constraints are defined by other, lower-level optimization and/or equilibrium problems that are parametrized by the variables of the higher-level problem. Problems of this type are difficult to analyze and solve, not only because of their size and complexity but also because they often fail to satisfy such standard assumptions as constraint qualifications and nondegeneracy.

With a focus on various aspects of HO problems with practical importance, the papers in this special issue of the journal Mathematical Programming, Series B are contributed by friends and admirers of the late Olvi L. Mangasarian, the John von Neumann Professor Emeritus of Computer Sciences, who passed away at the University of Wisconsin Hospital in Madison on March 15, 2020. The co-editors, all former students and colleagues of Olvi, are grateful to the authors and referees for their support of this special issue that we dedicate to the memory of Olvi's pioneering and visionary contributions to the field of optimization.

✉ Mikhail V. Solodov
   solodov@impa.br

   Kristin P. Bennett
   bennek@rpi.edu

   Michael C. Ferris
   ferris@cs.wisc.edu

   Jong-Shi Pang
   jongship@usc.edu

   Stephen J. Wright
   swright@cs.wisc.edu

[1]   Rensselaer Polytechnic Institute, Troy, USA

[2]   University of Wisconsin, Madison, USA

[3]   University of Southern California, Los Angeles, USA

[4]   IMPA, Rio de Janeiro, Brazil

The papers in this special issue cover a wide spectrum of theoretical and computational aspects of HO and related topics. They encompass deterministic and stochastic optimization problems with complementarity or equilibrium constraints, generalized Nash equilibrium problems, bilevel programming, and applications to statistics and machine learning.

Burtscheidt et al. consider a class of pessimistic bilevel stochastic optimization problems, and prove existence of solutions. The approach is applied to elastic shape optimization, demonstrating the interplay of follower and leader in design and testing.

Cui, Shanbhag, and Yousefian present complexity guarantees for an implicit smoothing-enabled method for stochastic optimization problems with equilibrium constraints. Their results cover certain classes of both single-stage and two-stage stochastic problems, as well as accelerated versions of algorithms.

Cui and Shanbhag consider computation of equilibria in monotone and potential stochastic hierarchical games. In this work, the game is noncooperative and each player solves a parametrized mathematical program with equilibrium constraints. For the monotone case, a variance-reduced stochastic proximal-point scheme is developed. When the game admits a potential function, the authors propose an asynchronous relaxed inexact smoothed proximal best-response framework.

Gahururo, Hintermüller, and Surowiec introduce risk-neutral PDE-constrained generalized Nash equilibrium problems, in which the feasible strategy set of each player is subject to a common linear elliptic partial differential equation with random inputs. Existence of equilibria and first-order optimality conditions are derived, and a relaxation scheme based on the Moreau-Yosida approximation of the bound constraint is proposed.

Gomez, He, and Pang study the single-parameter sparse optimization problem in statistical estimation defined by a pairwise separation objective. They introduce a linear-step inner-outer loop algorithm for computing a directional stationary solution of the nonconvex nondifferentiable folded concave sparsity problem. The authors also consider the parametric version of the problem that has a weighted l_1-regularizer and a quadratic loss function, and present a linear-step algorithm in two cases depending on whether the variables have prescribed or unknown signs.

Helou, Santos, and Simões propose a new approximate one-level reformulation for bilevel optimization problems. The reformulation is of the primal type, but is nonsmooth. An algorithm is developed that is proven to converge to a solution of the bilevel problem. It is shown that while each iteration depends on the solution of a nonsmooth optimization problem, it can be implemented in a computationally practical way.

Ho-Nguyen and Wright study a model for adversarial classification based on distributionally robust chance constraints. The authors show that under Wasserstein ambiguity, the model aims to minimize the conditional value-at-risk of the distance to misclassification. Also, a reformulation of the distributionally robust model for linear classification is provided, which is shown to be equivalent to minimizing a regularized ramp loss objective. Moreover, for a certain interesting class of distributions, the only stationary point of the regularized ramp loss minimization problem is the global minimizer.

Jiang and Chen formulate pure characteristics demand models under uncertainties of probability distributions as distributionally robust mathematical programs with stochastic complementarity constraints. To deal with uncertainties of probability distributions of the involved random variables in the second stage, the distributionally robust approach is employed. An approximating problem with regularization and discretization is proposed, which is a two-stage nonconvex-nonconcave minimax optimization problem. Convergence of the approximating scheme in the sense of the optimal solution sets, optimal values and stationary points is established.

Nie and Tang study convex generalized Nash equilibrium problems with polynomial data. Rational and parametric expressions for Lagrange multipliers are used to formulate polynomial optimization problems for efficient computation of equilibria via certain semidefinite relaxations. The method finds an equilibrium if one exists, or detects the nonexistence otherwise.

Shen, Ho-Nguyen, and Kılınç-Karzan propose an online framework for solving the convex bilevel optimization problem, where one optimizes a convex objective over the optimum solutions of another convex optimization problem. In this scheme, an online problem is formed in which the objective function remains fixed while the domain changes over time. Complexity guarantees are provided, as well as numerical illustrations on linear inverse problems and a large-scale text classification problem.

Ye, Yuan, Zeng, and Zhang present the difference-of-convex algorithms for solving bilevel programs in which the upper-level objective functions are differences of convex functions, and the lower-level programs are fully convex. Some numerical experiments for hyperparameter selection in machine learning illustrate the approach.

The above works owe much to Olvi's voluminous contributions, which, prior to the year 2000, are summarized in the guest editorial of *Computational Optimization and Applications*, Volumes 12 and 13 (1999) on the occasion of his 65th birthday. Since that time, Olvi made many further contributions to optimization and its applications, and in particular to optimization methods for machine learning. He had pioneered the latter area in a 1965 paper on pattern separation using linear and quadratic programming. During the past twenty years, the connections between optimization and machine learning have driven an enormous amount of research in both fields. Olvi's prescient early work predated this outburst of research activity by decades. We trust that the papers in this issue will spark new interest and activities in the area of hierarchical optimization, and build on some of the ideas that are pioneered in Olvi's vast research output.