PREFACE

Preface for the special issue of ADAC on 'Optimisation and Non-Convex Programming in Data Mining'

Hoai An Le Thi · Tao Pham Dinh · Gunter Ritter

Mathematical Programming, Operations Research, and Data Mining are closely related to each other by the fact that all three disciplines are, to a large extent, concerned with optimization problems. Moreover, all are directly oriented to practical problems that stem from the scientific, administrational, and industrial world.

As a major scientific challenge, Data Mining exploits large amounts of—in general multidimensional—data or other, possibly more complex information. Whereas the current computer technology enables us to collect and store large and even giant amounts of data it will still be necessary to develop new conceptual models, algorithms, and theoretical results in order to process and analyze such data and to find optimum constellations, parameters, classifications, etc. in an efficient way. Any success in this domain will entail direct consequences for many current and exciting research areas such as web page clustering, computer vision, financial mathematics, and bioinformatics. Perceptron, support vector machine, margin classifier, k-means clustering, and EM-algorithm are just a few keywords encountered in classification, clustering, and machine learning. Optimization is at the heart of such methodologies and therefore undoubtedly a modern key technology for Data Mining.

H. A. Le Thi (✉)
LITA, University of Paul Verlaine, Metz, Ile du Saulcy, 57045 Metz, France
e-mail: lethi@univ-metz.fr

T. Pham Dinh
LMI, National Institute for Applied Sciences, Rouen, BP 08, Place Emile Blondel,
76131 Mont Saint Aignan Cedex, France
e-mail: pham@insa-rouen.fr

G. Ritter
Faculty of Mathematics and Informatics, University of Passau, 94030 Passau, Germany
e-mail: ritter@fmi.uni-passau.de

This Special Issue of the journal *Advances in Data Analysis and Classification* collects a range of high quality research papers from the field of Optimization and Mathematical Programming and its application to Data Mining problems. More specifically, the four papers deal with the following topics:

- *Le Thi, Le Hoai, Nguyen, and Pham Dinh* are training a 0-norm Support Vector Machine after approximating the target function by the difference of two convex functions. The method allows to a certain extent the application of the machinery known from convex optimization, such as the subdifferential and the conjugate function, in order to locate critical points of the target function.
- *Mierswa and Morik* propose an explicit feature space transformation or detection of latent variables in nonconvex optimization problem induced by non-positive semidefinite kernel learning in order to better solve the related nonconvex programs. This paper motivates the usage of appropriate feature spaces and discusses the possible consequences leading to non-convex optimization problems.
- *Tarel, Ieng and Charbonnier* consider robust estimation when regression models are applied to real data as image analysis. They investigate Simultaneous Robust Multiple Fitting (SRMF) algorithm, whose global convergence to a local minimum is proved using results from constrained optimization theory. A parametric pdf family called SEF is introduced to reduce the dependency of SRMF on the initial conditions and the well-known GNC heuristic is used for this algorithm to better escape to local minima.
- *Dembélé* uses the Fuzzy C-Means method for clustering *3-way* gene expression data via optimization of multiple objectives. A reformulation of the total criterion is used to obtain an expression which has fewer variables compared to the classic FCM criterion. This transformation allows the use of a direct global optimizer in contrast to the alternating search commonly used.

Note that short versions of these papers were presented during the 'International Conference on Nonconvex Programming (NCP07): Local and Global Approaches—Theory, Algorithms, and Applications', held in Rouen/France from 17–21 December 2007.

Last but not least, this special issue would not have become possible without the encouragement and support by Hans-Hermann Bock, Editor in Chief of ADAC. Cordial gratitude to him by NCP07 and by the guest editors!