Math. Program., Ser. B (2011) 127:1–2
DOI 10.1007/s10107-010-0424-0

PREFACE

Preface

Alexandre d'Aspremont · Francis Bach ·
Inderjit S. Dhillon · Bin Yu

Published online: 23 October 2010

Statistics and optimization have long shared a common library of problems, yet it is only recently that significant advances in computing power have allowed mathematical programming to start attacking realistically large statistical problems, and statisticians to consider sophisticated optimization algorithms. This means that relatively advanced optimization results, e.g. in conic programming, duality theory, sensitivity analysis, complexity theory, now regularly appear in contributions to major machine learning conferences such as NIPS or ICML. At the same time, machine learning provides optimization with an ever larger array of new problems and challenging data sets: ℓ_1 penalized least-squares and the NETFLIX problem being two recent examples. In this spirit, this special issue mixes contributions from members of both communities.

Shai Shalev-Shwartz, Yoram Singer, Nathan Srebro and Andrew Cotter, discuss their recently introduced and very popular training algorithm for support vector machines, in a paper entitled "Pegasos: Primal Estimated sub-GrAdient SOlver for SVM".

A. d'Aspremont (✉)
ORFE Department, Princeton University, Princeton, NJ 08544, USA
e-mail: aspremon@Princeton.edu

F. Bach
INRIA - Willow project, Laboratoire d'Informatique de l'Ecole Normale Supérieure,
Paris, France

I. S. Dhillon
Department of Computer Sciences, University of Texas at Austin, Austin, TX, USA

B. Yu
Department of Statistics, Department of Electrical Engineering & Computer Science,
UC Berkeley, Berkeley, CA, USA

First-order methods are often the only option when solving large-scale machine learning problems and Yurii Nestrov presents here a "Barrier subgradient method" and derives a new affine-invariant primal-dual subgradient method for nonsmooth convex optimization problems, where the feasible set is described by a self-concordant barrier. The paper also begins with a high-level classification of current complexity results based on the tractability of elementary operations on the problem (e.g. maximizing affine functions on the feasible set, computing derivatives, etc.).

Three papers in this issue focus on the problem of efficiently bounding the performance of ℓ_1 decoding based on properties of the nullspace of the coding matrix. In their paper "On Verifiable Sufficient Conditions for Sparse Signal Recovery via ℓ_1 Minimization", Anatoli Juditsky and Arkadi Nemirovski produce linear programming based relaxations for the problem of computing sparse recovery thresholds and tightly bound the performance of their result. In "Verifiable conditions of ℓ_1-recovery for sparse signals with sign restrictions", they extend these results to the case where the variables are constrained to be nonnegative. Alexandre d'Aspremont, Laurent El Ghaoui also produce performance bounds for sparse recovery using semidefinite relaxation techniques in a paper entitled "Testing the Nullspace Property using Semidefinite Programming."

In a paper on "Chance Constrained Uncertain Classification via Robust Optimization", Aharon Ben-Tal, Sahely Bhadra, Chiranjib Bhattacharyya and J. Saketha Nath relax a chance constrained programming formulation of the support vector machine problem with uncertain training data.

Finally, two papers consider two facets of recommendation systems: Benjamin Recht, Weiyu Xu and Babak Hassibi study explicit recovery conditions on the (related) problem of matrix completion from a few entries in a paper entitled "Null Space Conditions and Thresholds for Rank Minimization". Xiaoye Jiang, Lek-Heng Lim, Yuan Yao and Yinyu Ye show how pairwise rankings may or may not be turned into a global ranking, with explicit notions of local and global inconsistencies, in a paper entitled "Statistical Ranking and Combinatorial Hodge Theory".