REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
30-04-2012

2. REPORT TYPE
Final Report

3. DATES COVERED *(From - To)*
20090315-20111130

4. TITLE AND SUBTITLE
Discovering Structure via Matrix Rank Minimization

5a. CONTRACT NUMBER
FA 9550-09-1-0247

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Vavasis, Stephen A.
Wolkowicz, Henry

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Waterloo
200 University Ave W.
Waterloo, Ontario, Canada N2L 3G1

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Science and Research
875 Randolph Street
Suite 325 Room 3112
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)
AFOSR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-OSR-VA-TR-2012-0968

12. DISTRIBUTION/AVAILABILITY STATEMENT
Public distribution

Distribution A: Approved for Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT

During this three-year project, Vavasis and Wolkowicz, together with students, postdocs, and colleagues, made a number of advances in the convex optimization and its application to data mining and sensor localization. In this report we highlight some of these accomplishments.

15. SUBJECT TERMS
Optimization, convexity, relaxation, matrix rank minimization, NP-hard problem, data mining, clique, biclique, nonnegative matrix factorization, sensor localization, distance geometry

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Stephen A. Vavasis
UU	UU	UU	UU	7	19b. TELEPHONE NUMBER *(Include area code)* 519-888-4567 ext 32130

Discovering structure via matrix rank minimization: Final Report

S. Vavasis H. Wolkowicz

2012-April-25

1 Overview

During this three-year project, Vavasis and Wolkowicz, together with students, postdocs, and colleagues, made a number of advances in the convex optimization and its application to data mining and sensor localization. In this report we highlight some of these accomplishments.

2 Research progress in data mining

During this project, we made progress on several fronts with regard to convex relaxation of problems in data mining and sensor network location. Ames and Vavasis completed two papers on data mining with convex optimization. The first, "Nuclear norm minimization for the planted clique and biclique problems," considers the following NP-hard problems. The *maximum clique* problem is, given a graph G, find the largest set of nodes that are all mutually interconnected by edges , i.e., a set of k nodes such that all possible $k(k+1)/2$ edges are present between the nodes, with k maximum. The *maximum biclique* problem is, given a bipartite graph $G=(U,V,E)$, find the $U^* \subset U$ and $V^* \subset V$ such that all possible edges $|U^*| \times |V^*|$ are present in the graph, and such that $|U^*| \cdot |V^*|$ is maximized. Both problems are NP-hard.

Both problems are simple model prolems of information retrieval problems. For example, the clique problem has been used on EEG data to find sites in the brain linked to seizures, while the biclique problem has been used to find features in databases of images.

Ames and Vavasis [2] showed that if the clique or biclique is constructed in the following manner, then it is solvable in polynomial time via convex relaxation. Start with a graph containing n nodes and no edges. Insert either a single clique or biclique, say with k nodes and $k(k+1)/2$ edges (in the clique case) or $k+k'$ nodes and kk' edges (in the biclique case). Then inserted diversionary edges either at random or according to an adversary's strategy. If the edges are inserted at random, then the graph size n may be as large as const $\cdot\, k^2$ (and hence with constk^4 edges), and yet the clique is still found via convex relaxation with probability exponentially close to 1. If the edges are inserted by an adversary, then convex relaxation can still find the clique or biclique even though as many as const $\cdot\, k^2$ have been inserted (almost enough to create another maximum clique or biclique).

Their second paper concerned the combinatorial clustering problem [3]. This problem asks, given a graph G in which nodes are data atoms and edges are affinities and given an integer k, find k cliques in the graph that cover the maximum number of nodes. This problem is obviously NP-hard since it generalizes the famous max-clique problem (the $k = 1$ case). It is also a very general way to pose clustering problems which try to partition data items into clusters so that the affinities between atoms in the same cluster are greater than affinities between atoms in other clusters. The result of Ames and Vavasis is that this NP-hard problem can be solved in polynomial time by a convex relaxation provided the instance is constructed in a certain way. The construction involves first laying down the clusters and then additional inserting a large amount of noise, that is, affinity edges between nonrelated nodes chosen at random.

The third paper coming out of Ames' PhD thesis is singly authored by him. In this paper, he considers a more general version of clustering in which edges between nodes in the same cluster are assigned a weight drawn from a probability distribution with mean α, while nodes in different clusters or nodes not in any cluster are drawn from a probability distribution with mean β. Surprisingly, the assumption $\alpha > \beta$ is the only one needed to ensure that a convex relaxation can recover the clusters provided they satisfy a certain size lower bound. This is a guarantee stronger than any we know about in the previous literature.

We also carried out work with Xuan Vinh Doan, a postdoc hired under the project, on feature extraction from data matrices. Doan and Vavasis's first result addresses the NP-hard problem of finding a large approximately rank-one submatrix of a given nonnegative matrix A. If A encodes an image database, then finding this submatrix can help find a visual feature in an image database or a topic in a text database. We are able to solve this problem in polynomial time using convex relaxation, provided the input data matrix is constructed in a certain way. This result has been submitted to a journal [9]. Our second result addresses efficient solution of the resulting relaxation [8]. Our goal is to develop an algorithm for nonnegative matrix factorization (NMF), that is, a decomposition of A into nonnegative rank-one terms. This problem is NP-hard but may admit a solution via convex relaxation that is guaranteed to work for certain classes of instances. We have begun work with V. Chandrasekaran on convex relaxation of NMF.

3 Research progress in convex optimization

Karimi and Vavasis have developed a new algorithm, CGSO [12], for large scale unconstrained convex optimization in the case that the objective function is differentiable. Their method combines the best features of classical nonlinear CG with the complexity bounds of Nemirovsky and Yudin. In more detail, classical nonlinear CG (e.g., Fletcher-Reeves, Polak-Ribière) are generalizations of Hestenes-Stiefel conjugate gradient and, as such, converge in an optimal fashion for convex quadratic objective functions. On the other hand, for general differentiable objective functions their performance can be exceedingly poor as demonstrated by Nemirovsky and Yudin. On the other hand, Nemirovsky and Yudin's method has the optimal complexity bound up to a constant factor, but is not exactly optimal, and in particular, is suboptimal for quadratic functions. In addition, their algorithm requires knowledge of Lipschitz constants of the function.

Our proposed algorithm is optimal for quadratic functions (because it reduces to H-S) but also achieves the Nemirovsky-Yudin bound for general differentiable convex functions. In addition, it does not require knowledge of any parameters. Each iteration requires solution of a subproblem in two dimensions, which is more expensive than traditional CG. In addition, if certain inequalities fail, the dimension of the subproblem can increase to up to $O(\log k)$, but we have not seen this behavior in practice. Even though the subproblem is more expensive, our computational testing shows that the method outperforms traditional conjugate gradient in almost all cases.

In current work begun at the end of the grant, Karimi and Vavasis are extending the method to constrained convex optimization. A targeted application is solving the convex relaxation arising in compressive sensing.

We report on our progress in developing efficient, stable, algorithms for three classes of problems: Linear Programming (LP), Semidefinite Programming (SDP), Trust Region Unconstrained Minimization (TR). Wolkowicz and students and postdocs have developed an approach to making interior point methods more robust. Current popular interior-point path-following algorithms for LP and SDP use block elimination to exploit the structure of the optimality conditions and form the *normal equations* to find the Newton search direction. The block eliminations do not include any type of *partial pivoting*. This leads to instability and ill-conditioning. Their approach avoids the ill-conditioning to obtain a stable search direction that can exploit sparsity. This search direction allows for solutions with higher accuracy. The LP code from [11] was recently revised and implemented into the latest version `http://www.maplesoft.com/`MAPLE 14 (work with Jason Hinek). The SDP code has been written and rigorously tested in [10] and is being implemented for the next version of MAPLE. In addition, an efficient TR method based on a trust region Lanzcos type subproblem algorithm that can exploit sparsity. (This is part of the Master's thesis of Heng Ye.)

4 Research progress in sensor localization and Euclidean distance matrices

There are many instances of SDP where the Slater constraint qualification (CQ) fails. A facial reduction process can be used to regularize the problem. This is studied and implemented in [5]. In particular, the CQ fails for instances of sensor network localization (SNL) and molecular conformation (MC). Instead of this being a drawback, the structure can be exploited to both speed up algorithms and increase the accuracy. For SNL, the facial reduction corresponds to exploiting the clique structure of the problem. This has been implemented and tested to solve huge problems to high accuracy, e.g. $n = 10^6$ nodes to 16 decimals accuracy in under 3 minutes on a laptop, see [14] as well as e.g. [15, 13, 1].

For MC, typical proteins are formed from amino acids whose molecular structure is well known. This is equivalent to fixing cliques in the protein and applying facial reduction. This research is ongoing with several publications in progress, and a poster session at `http://compbio.cs.sfu.ca/recomb2011/`RECOMB 2011 (work with Nathan Krislock, Babak Alipanahi, Ming Li, et al). The facial reduction idea has been successful in increasing the accuracy and reducing the computation time. The technique has been implemented in a software package called SPROS, which is able to determine protein structures as accurately as any known competing approach. Furthermore, SPROS is more robust with respect to noise in the data than other methods because the noise can be properly accounted for via inequality constraints rather than ad-hoc techniques in which some of the distances are assumed to be exact.

5 Outreach activities

Vavasis served on the program committee of the SIAM Annual Meeting, Pittsburgh, PA, July 12–16, 2010. He also organized two minisymposia on topics related to the project, namely, first order methods for convex optimization and recent progress in rank minimization. The eight speakers in these two minisymposia were S. Wright (Wisconsin), S. Karimi (Waterloo), G. Lan (Florida), Z. Lu (Simon Fraser), K.-C. Toh (National U. Singapore), I. Dhillon (Texas), B. Ames (Waterloo), M. Fazel (Washington). Finally, he served as a panelist on a career development panel.

Wolkowicz served on the ICIAM 2011 Program Committee and organized three minisymposia. He also served as editor of a special issue of Math. Prog. He also served on the organizing committee for the workshop in honor of J. Borwein's 60th birthday and as an editor of the proceedings.

Vavasis also served as co-chair of the SIAM Conference on Optimization, Darmstadt, Germany, May 16–19, 2011. In addition to inviting speakers, organizing sessions and other administrative work, he also organized and chaired the funding agency panel.

Vavasis's students and postdocs gave talks on work related to the project at the Fall 2011 INFORMS meeting, the SIAM Annual Meeting, and the Waterloo Grad Student Research Conference. Wolkowicz delivered a plenary talk at the workshop in honor of Claude Lemarechal.

Vavasis served as Program Director for the SIAM Activity Group in Optimization and as a member of the Householder Prize Committee. Wolkowicz served as a member of the SIAM Council.

The grant hosted visits by Serge Kruk (Oakland), Boris Mordukhovich (Wayne State), Kim-Chuan Toh (National U. Singapore), Yurii Nesterov (U. C. Louvaine), Inderjit Dhillon (Texas), Yuri Levin (Queens), Yinyu Ye (Stanford), Pablo Parrilo (MIT), Alexander Barvinok (Michigan).

References

[1] A. Alfakih, M.F. Anjos, V. Piccialli, and H. Wolkowicz. Euclidean distance matrices, semidefinite programming, and sensor network localization. *Portug. Math.*, to appear.

[2] B. Ames and S. Vavasis. Nuclear norm minimization for the planted clique and biclique problems. *Mathematical Programming*, 129(1):69–89, 2011.

[3] Brendan P.W. Ames and Stephen A. Vavasis. Convex optimization for the planted k-disjoint-clique problem. Submitted to Math. Prog., 2010.

[4] A. Babak, N. Krislock, A. Ghodsi, H. Wolkowicz, L. Donaldson, and M. Li. Spros: An sdp-based protein structure determination from nmr data. Technical report, University of Waterloo, Waterloo, Ontario, 2011. poster session at RECOMB2011.

[5] Y-L. Cheung, S. Schurr, and H. Wolkowicz. Preprocessing and reduction for degenerate semidefinite programs. Technical Report CORR 2011-02, University of Waterloo, Waterloo, Ontario, 2011.

[6] Y. Ding, D. GE, and H. Wolkowicz. On equivalence of semidefinite relaxations for quadratic matrix programming. *Math. Oper. Res.*, 36(1):88–104, 2011.

[7] Y. Ding, N. Krislock, J. Qian, and H. Wolkowicz. Sensor network localization, Euclidean distance matrix completions, and graph realization. *Optim. Eng.*, 11(1):45–66, 2010.

[8] X. V. Doan, K.-C. Toh, and S. Vavasis. A proximal point algorithm for finding approximately rank-one submatrix with nuclear norm and ℓ_1 norm. Submitted, SIAM J. Sci. Comput., 2011.

[9] X. V. Doan and S. Vavasis. Finding approximately rank-one submatrices with the nuclear norm and ℓ_1 norm. In review process, SIAM J. Optimiz.; URL:`http://arxiv.org/abs/1011.1839`, 2010.

[10] X.V. Doan, S. Kruk, and H. Wolkowicz. A robust algorithm for semidefinite programming. Technical Report CORR 2010-09, University of Waterloo, Waterloo, Ontario, 2010. submitted in November, 2010.

[11] M. Gonzalez-Lima, H. Wei, and H. Wolkowicz. A stable primal-dual approach for linear programming under nondegeneracy assumptions. *Comput. Optim. Appl.*, 44(2):213–247, 2009.

[12] S. Karimi and S. Vavasis. Conjugate gradient with subspace optimization. Draft manuscript, 2011.

[13] N. Krislock, F. Rendl, and H. Wolkowicz. Noisy sensor network localization using semidefinite representations and facial reduction. Technical Report CORR 2010-01, University of Waterloo, Waterloo, Ontario, 2010.

[14] N. Krislock and H. Wolkowicz. Explicit sensor network localization using semidefinite representations and facial reductions. *SIAM Journal on Optimization*, 20(5):2679–2708, 2010.

[15] N. Krislock and H. Wolkowicz. Euclidean distance matrices and applications. In *Handbook of Semidefinite, Cone and Polynomial Optimization: Theory, Algorithms, Software and Applications*, number 2009-06 in CORR. Springer-Verlag, Waterloo, Ontario, to appear.

[16] H. Wei and H. Wolkowicz. Generating and solving hard instances in semidefinite programming. *Math. Programming*, 125(1):31–45, 2010.

[17] H. Wolkowicz. Generating eigenvalue bounds using optimization. In *Nonlinear analysis and variational problems*, volume 35 of *Springer Optim. Appl.*, pages 465–490. Springer, New York, 2010.