Convex relaxations of NP-hard problems
and
Efficient Robust Financial Optimization

S. Vavasis and H. Wolkowicz

September 26, 2009

1 Convex relaxations of NP-hard problems

Vavasis and Wolkowicz will develop new results and applications for matrix rank minimization problems. *Matrix rank minimization* refers to the optimization problem of minimizing the rank of an unknown matrix X subject to the constraint that X lies in a a certain specified convex set C. This problem is NP-hard even when the set C is an affine linear set, i.e., when X is constrained to satisfy a system of linear equalities. Despite the NP-hardness of the problem, it has emerged recently that interesting instances of matrix rank minimization can be solved, sometimes to optimality, in polynomial time. Furthermore, many useful applications of matrix rank minimization have appeared in the literature, some of which are described below, and therefore the problem merits more intensive study.

Let us first consider linear equality constrained matrix rank minimization, that is, the problem of minimizing the rank of X subject to the constraint $\mathcal{A}X = b$, where $X \in \mathbb{R}^{m\times n}$ is the unknown matrix, $\mathcal{A} :\to \mathbb{R}^p$ is a linear transformation that maps X to a p-vector, and b is a p-vector. This problem is known to be NP-hard. In particular, in the special case that X is constrained to be diagonal, this problem specializes to the "sparsest vector" problem (find the vector in an affine space with the fewest number of nonzero entries). The sparsest vector problem is NP-complete, a result known since the 1970s.

In 2007, Recht, Fazel and Parrilo [1] discovered the following intriguing result. If $m, n, p, \mathcal{A}$ are as given above where $\mathcal{A}$ is a randomly chosen linear transformation (chosen according to a particular distribution that is quite reasonable), and if there is a sufficiently low-rank solution X^* to the problem, then this optimal solution X^* can be discovered by minimizing $\|X\|_*$ subject to $AX = b$. Here $\|\cdot\|_*$ denotes the *nuclear norm*, that is, the sum of the singular values of X. This nuclear norm minimization optimization is convex and indeed can be reformulated as a semidefinite programming problem. Therefore, the nuclear norm relaxation can be solved to any desired accuracy in polynomial time.

This result is closely related to compressive sensing, a topic that burst into the literature in about 2005, with the publication of several hundred papers by now. The main result of compressive sensing is that if A is a randomly chosen $m \times n$ matrix with $m \ll n$ and $\mathbf{x}$ is a sufficiently sparse vector (in particular, $\mathbf{x}$ has fewer than m nonzero entries), then $\mathbf{x}$ can be recovered from knowledge of $b = A\mathbf{x}$ using convex optimization. Thus, this result shows that the NP-hard *sparsest vector* problem can be solved in polynomial time provided the instance is constructed in this manner. Compressive sensing is expected to have widespread applications in signal and image processing. Indeed, there are plans to build efficient digital cameras based on compressive sensing.

The Recht, Fazel and Parrilo result is quite intriguing but does not seem to be widely applicable because rank minimization problems that occur naturally do not typically have constraints of the form $\mathcal{A}X = b$ with a random transformation $\mathcal{A}$. Other recent works consider rank minimization with more natural constraints. For example, Candés and Recht and later Candés and Tao have considered the *matrix completion problem*, which is the problem of minimizing the rank of a matrix B subject to the constraint that certain entries of B are specified. They show that if the known entries follow a pattern that is random with respect to the singular vectors of the unknown matrix, then the nuclear norm relaxation is exact. Another matrix completion problem is the sensor localization problem, which we consider below.

A second recent rank minimization paper is from our group (Ames and Vavasis) and considers the maximum clique and biclique problems. The *maximum clique* problem is as follows. One is given an undirected graph and asked to find the largest k-clique, that is, a subgraph of k completely interconnected nodes. The *maximum biclique* problem asks: given a bipartite graph G, find the subgraph that contains m vertices of the left node set, n of the right node set, and all possible connecting mn edges, such that mn is maximized. Both problems are NP-hard.

Ames and Vavasis rewrite both clique and biclique as matrix rank minimization. Then they pass to the nuclear norm relaxation. They prove that the nuclear norm relaxation can find the maximum clique or biclique of the original instance for classes of instances that comprise a single clique or biclique plus a number of diversionary edges inserted into the graph. They consider two ways to insert diversionary edges, either by an adversary who can add the diversionary edges in an arbitrary manner, or by a random process that selects each graph edge outside the clique to be diversionary with a certain fixed probability. In the case of the maximum clique, they prove that the nuclear norm relaxation can find a clique of size n provided that the adversary can insert no more than $O(n^2)$ edges and, each nonclique node is adjacent to no more than $O(n)$ clique nodes. In the case of random diversionary edges, they prove that a clique of size n embedded (hidden) in a random graph with $O(n^2)$ edges can still be found with the nuclear norm relaxation. Analogous results are established for the biclique problem.

The clique and biclique problems were chosen for this study because they both arise in data mining. In particular, the clique problem has been used in the study of brain activity leading up to seizures in epileptic patients, and the biclique problem has been used to find features in image databases. For any kind of data mining problem, a reasonable question to ask is whether an algorithm is able to find the hidden structure under the assumption that it is present but is perhaps obscured by noise. Our answer is affirmative for the nuclear norm algorithm for both clique and biclique under suitable assumptions.

In the new term of the MITACS grant, we will investigate the use of nuclear norm and other convex relaxations for further data mining problems. Three data mining problems of interest to us include nonnegative matrix factorization, clustering, and tensor decomposition. Nonnegative matrix factorization (NMF) is used to find features in image, text and biochemical experimental databases. We have already completed two papers recently with novel results about NMF. In one paper, we proved that NMF is an NP-hard problem. In the other, we proposed a new greedy heuristic for the problem that comes with some theoretical guarantees. We believe that a convex relaxation of this problem will yield a better factorization (and hence be able to find features more effectively) than any existing method. Clustering refers to the problem of automatically partitioning entries in a database into subsets so that the members of each subset are closer to each other than to the other subsets. Finally, tensor decomposition refers to the problem of factorizing an order 3 or higher tensor. Many of the entries may be unknown. We have begun preliminary discussion with Ontario IESO (operator of the Ontario electric power grid) about a project to develop a simple model of a customer demand function for electricity by applying tensor factorization to a database of historical price information.

Another important application of matrix completion is the sensor localization problem, SNL. The problem here is to find the locations (in embedding dimension $r = 2$, (x, y), or $r = 3$, (x, y, z) coordinates) of n sensors that form an intercommunicating network. The given data consists of the locations of a few of the sensors (the *anchors*) as well as the Euclidean distance between certain pairs of sensors, usually those that are within radio range of each other. The problem in general is also NP-hard. It can also be framed naturally as rank minimization. In particular, let D be the *Euclidean distance matrix*, that is, the matrix whose (i, j) entry is the squared distance $\|\mathbf{x}_i - \mathbf{x}_j\|^2$, where for each i, $\mathbf{x}_i$ denotes the location of the ith sensor. Let $P = \begin{bmatrix} X \\ A \end{bmatrix}$, where the rows of X and A denote the positions of the sensors and anchors, respectively. Then there is a linear relationship between the positive semidefinite matrix $B = PP^T$ and D, denoted $\mathcal{K}(B) = D$. The EDM completion problem can now be solved by finding the positive semidefinite matrix b with the correct rank that corresponds to the given embedding dimension ($r = 2$ or $r = 3$). The linear constraints are given by $(\mathcal{K}(B))_{ij} = D_{ij}$, for all known distances D_{ij}.

Recent results in our group shows that the SNL problem is implicitly highly degenerate. But, one can take advantage of the degeneracy to develop an algorithm called *facial reduction* that is often able to solve huge instances of the sensor localization problem very efficiently. For example, if a block of distances $D[\alpha]$ is known, for some index set $\alpha \subset \{1, \ldots, n\}$, with $|\alpha| = k$, then the rank of the Moore-Penrose inverse $\mathcal{K}^\dagger(D[\alpha])$ is r. This restricts the the rank of the corresponding block in B to be at most $r + 1$. Therefore, the rank of B is restricted by $n - k + r$, i.e. the standard Slater constraint qualification fails. The algorithm then finds the correct face of the semidefinite cone to reduce the problem. This information can be used to iteratively collapse the problem to a very small semidefinite programming problem. Problems with up to 10^5 sensors are solved in minutes on an ordinary laptop.

In general, the convex relaxation of the SNL problem does not find a sufficiently low rank positive semidefinite B that satisfies all the equality constraints. ???So???or?? Biswas?? and Ye cite??? propose a more sophisticated relaxation that has guarantees of finding the low-rank solution under sufficient assumptions.

We propose to develop new approaches for sensor localization based on our own facial reduction combined with Candès-Tao. Candès-Tao covers sensor localization in the case that the known distances are randomly chosen (an unrealistic model), whereas facial reduction relies on the fact that in typical networks, there are paths made of small cliques in the distance data. On the other hand, we do not know of a useful distribution of distance data that is known with high probability to work for the facial reduction algorithm. The question is then whether there exists a more realistic but still randomized model that is guaranteed to be solved with facial reduction, or perhaps facial reduction combined with the solution of small (local) convex programming problems.

A final issue of interest to us is the following. In the results mentioned so far about nuclear norm relaxation, the problem is solved all at once with a single convex programming problem. What if one is allowed to solve a polynomial number of convex programming problems? Does this substantially enlarge the class of instances that are solvable in polynomial time? ????another final issue???? what happens with noisy data??? makes the problem much harder????add a comment???

2 Efficient Solutions for Robust Financial Optimization Problems

Although financial optimization has been an area of active research for the past forty years, dealing with uncertainty in parameter estimation remains among the key issues that are still unresolved. For several decades, stochastic programming was one of the major techniques used to address these problems, but recent advances in robust optimization have opened up new approaches for modeling uncertainty and thus offering new opportunities in the emerging area of Robust Financial Optimization (RFO).

This novel framework captures the uncertainty in a generic way without increasing the complexity of the original deterministic model and produces computationally tractable formulations without the inherent limit to low dimensions of classical approaches. The novel RFO framework avoids some of the difficulties of the stochastic-based models, where the size of the problem grows exponentially with the number of uncertainty sources and the span of time horizon. It also relaxes the restrictive assumption that probability distributions need to be known a priori. Many robust optimization models can be represented as a linear, second-order conic, or semi-definite programming problems, so that IPMs and other algorithms for conic optimization can be applied to solve instances of these models.

Our project plans on extending our work on interior-point methods for linear, semidefinite, and second-order cone programming to the special structures that arise in RFO. We aim to develop specialized algorithms and software tools that enables the solution of large-scale RFO models within the required time frame.

In addition, we intend to establish the connection between risk measures and robust uncertainty sets and develop better models for portfolio optimization, Sharpe ratio maximization, and asset liability management in multi-stage and uncertain environments in the presence of realistic constraints and assumptions. Further, we aim to develop novel methods and techniques that, based on novel parametric programming methods, enables us to produce the efficient (Pareto) frontier in case multiple, conflicting objectives are present. Finally, we plan on implementing efficient algorithms for industrial large-scale problems.

References

[1] B. RECHT, M. FAZEL, and P. PARRILO. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. Technical report, Caltech, Pasadena, California, 2007.