Final Scientific/Technical Report

Project: Algorithms for Mathematical Programming with Emphasis on Bi-level Models
Grant Number: DE-FG02-08ER25856
Period Covered: August 15, 2009 – August 14, 2013
PI: Donald Goldfarb (goldfarb@columbia.edu)
co-PI: Garud Iyengar (garud@ieor.columbia.edu)
PI Institution: Columbia University

The research supported by this grant was focussed primarily on first-order methods for solving large scale and structured convex optimization problems and convex relaxations of nonconvex problems. These include optimal gradient methods, operator and variable splitting methods, alternating direction augmented Lagrangian methods, and block coordinate descent methods. We now describe this research.

We have developed two new first-order methods for solving the basis pursuit (BP) problem $\min\{\|x\|_1 : Ax = b\}$. In [11] we describe a Sequence Penalty Algorithm (SPA) that computes a feasible solution for the BP problem by solving a sequence of penalty problems. This algorithm computes an ϵ-optimal and ϵ-feasible solution in $\mathcal{O}(\epsilon^{-\frac{3}{2}})$ iterations, where each iteration takes $\mathcal{O}(n\log(n))$ operations whenever the matrix A is a partial Fourier, DCT or Wavelet matrix. In particular, we do not require that A be an orthogonal matrix. The SPA can also efficiently solve problems of the form $\min\{\sum_j \|B_j x\|_1 : Ax = b\}$. Problems of this nature naturally appear whenever we require some linear function of x to be sparse, i.e we may require that x is piecewise flat, or that its derivative in a given direction is piecewise flat. In [12] we describe a First-order Augmented Lagrangian (FAL) algorithm that computes an ϵ-optimal and ϵ-feasible solution to the basis pursuit problem in $\mathcal{O}(\frac{1}{\epsilon})$-iteration. In each step of the algorithm we solve a constrained shrinkage problem. This algorithm performs very well in practice; in particular, it is able exactly recover the zero-set (i.e. the set of components where $x_i = 0$) without any post-processing. In [28] we develop an augmented Lagrangian technique with increasing penalty to solve a composite norm minimization problem, and in [31] we extend this technique to solving general conic convex programs. In [30], we propose an alternating direction method with increasing penalty for the stable principal component pursuit problem.

We have combined a first-order operator splitting method (fixed-point continuation (FPC)) with active set identification and subspace optimization to develop an extremely effective and efficient method [8] for the BP problem. This method is proved to converge R-linearly and globally in [18] and it has been shown to be able to recover signals with very large dynamic ranges in compressed sensing applications. We have also developed an FPC method and a Bregman iterative algorithm to solve matrix rank minimization problems [10,37] and analyzed their ability to recover low-rank solutions [15]. In addition, in [21] we have developed accelerated versions of the linearized Bregman method and proved that their iteration complexity is reduced from $\mathcal{O}(1/\epsilon)$ to $\mathcal{O}(1/\sqrt{\epsilon})$ to obtain an ϵ-optimal solution and applied them to compressed sensing and matrix completion problems.

In [20] we developed alternating linearization methods (ALMs) for solving convex optimization problems that often arise as tight convex relaxations of nonconvex structured optimization problems. Our methods solve problems of the form:

$$\min\{F(x, y) \equiv f(x) + g(y) : Ax + y = b\}.$$

Under the assumption that both f and g are convex functions with Lipschitz continuous gradients, we prove that our methods require $O(1/\epsilon)$ iterations to obtain an ϵ-optimal solution. We also propose accelerated versions of our methods that have an iteration complexity of $O(1/\sqrt{\epsilon})$, while requiring essentially the same computational effort at each iteration.

We developed specialized versions of these algorithms to solve various high dimensional problems in statistical learning. Specifically, we applied them to Gaussian graphical models (sparse inverse covariance selection) in [39] and to overlapping group LASSO problems involving appropriate sparsity-inducing norm regularizers in [17]. We developed an alternative efficient block-coordinate descent approach for solving group LASSO problems in [22]. We also, developed line search versions of our accelerated ALMs and the fast prox-gradient FISTA method that preserve these methods' fast iteration complexity in [25], and specialized versions of ALMs and prox-gradient methods for solving robust and stable principle component pursuit problems in [24].

In [19] we developed general K-splitting algorithms (K can be any finite number) for solving convex optimization problems with Lipschitz continuous gradients. These methods are Gauss-Jacobi-like and hence are parallelizable, which makes them particularly attractive for solving large-scale problems. We prove that the number of iterations needed by the first class of algorithms to obtain an ϵ-optimal solution is $O(1/\epsilon)$. The algorithms in the second class are accelerated versions of those in the first class, where the complexity result is improved to $O(1/\sqrt{\epsilon})$, while the computational effort required at each iteration is essentially unchanged. To the best of our knowledge, the complexity results for both our Gauss-Seidel and Jacobi-like methods are the first such results to have been given for multiple splitting and alternating direction type methods.

In [9] we develop a third class of methods that is based on alternating direction schemes for minimizing the dual augmented Lagrangian function for an SDP. For these methods we have only partial convergence results. However, numerical results that we have obtained for frequency assignment, maximum stable set and binary integer quadratic programming problems demonstrate the robustness and efficiency of this approach.

We also developed an overlapping block-coordinate descent method for solving SDP problems [43], based on relaxing the n-dimensional positive semidefinite constraint on the matrix X. By fixing any $(n-1)$-dimensional principal submatrix of X and using its (generalized) Schur complement, the positive semidefinite constraint is reduced to a simple second-order cone constraint. When this method is applied to solve the maxcut or matrix completion SDP relaxations, closed-form solutions for the subproblems are available. Our numerical results on large scale instances of these problems show that these methods are extremely fast.

We have also developed algorithms for de-noising images that use network flow approaches [3] that are based on binary Markov random field models and that use the Gallo-Grigoriadis and Tarjan's parametric version of the Goldberg-Tarjan preflow-push max-flow algorithm combined with a divide-and-conquer approach, ones that use an alternating direction augmented Lagrangian approach [35], and curvilinear search methods for color images [4]. Our most recent work on first-order algorithms has focused on tensor completion and and low-rank tensor recovery [27] and theoretical recovery guarantees that are somewhat analogous to those known for the corresponding matrix cases. Specifically, for recovering tensors with low Tucker rank, the matrix unfoldings of the tensor need to satisfy a mutual incoherence condition in addition to the usual incoherence conditions for each unfolding [29,36].

For general (nonconvex) nonlinear programming (NLP), we developed an algorithm that combines a piecewise linear penalty function approach with an interior-point ℓ_2-penalty approach. The resulting method [5] has properties of both a penalty method and a strengthened filter method, and in particular, very strong global and local convergence properties. Extensive computational tests show that our implementation is as effective as state-of-the-art NLP codes.

To solve large-scale, possibly non-convex optimization problems that arise from inverse problems and other infinite dimensional problems, we developed a globally convergent line search multigrid (multiscale) method in [2] that makes minimal demands on the minimization method used on each grid level.

We have continued to work on semidefinite packing and covering problems. In [14], we extended our previous work to a much larger class of semidefinite packing problems. This set of problems includes the sparse Principal Component Analysis (PCA) problem. We show that our proposed algorithm is faster than all previously known algorithms for this problem – both in theory and in practice. In [38] we developed algorithms for approximately solving semidefinite covering problems. This is a non-trivial extension that required the development of new methods of analysis. We also reported the performance of our implementation of the proposed algorithms. The main bottleneck in these algorithms is computing the exponential of a sparse matrix.

We continued working on applications of robust optimization in portfolio management and asset-liability management. In [7] we developed techniques for ensuring that a pension fund is meets its obligations when there is uncertainty in the equity return and the yield curve. Our model also allows us to predict the worst-case contribution corresponding to a fixed portfolio management scheme. In [1] we showed how to use robust optimization techniques to solve a cash flow problem when the yield curve is uncertain and there are several different analyst views on the relative movement of interest rates. The current state-of-the-art technique for solving these problem assumes that the yield curve is described a copula model and then incorporates the analysts view into the copula model using Bayes rule. This method does not scale very well as the time horizon as well as the number of bonds available to hedge a portfolio. Our model results in a linear program that can be solved very efficiently using dual-decomposition techniques. In [6], we developed a behavioral-finance based model for joint evolution of price and volumes. We developed a convex optimization based methodology to efficiently calibrate the model to observed data and predict prices and volumes in the future. In [16] and [32], we proposed a fast first-order algorithm for solving portfolio selection problems with multiple spectral risk constraints (a spectral risk function is a convex combination of conditional value-at-risk constraints). Portfolio selection with spectral risk constraints can be approximated by a linear program. The running time of our proposed algorithm is at least two orders of magnitude smaller than that of the state-of-the-art solver applied to the linear program. Moreover, unlike the LP-based approach, the running time of our proposed algorithm is robust with respect to perturbations to problem data.

In [13] we investigated extremal income inequality in a stable network. Here the "stability" is defined to mean that no subset of nodes of the network has an incentive to secede from the grand coalition of all node, and the income distributions are compared using the Lorenz ordering. Previously it was conjectured in the literature that the extremal income distribution is function of the degree distribution – graphs with more unequal degree distribution would support more income inequality. We showed that for bipartite graphs, the income inequality is completely determined by the size of the maximal independent set. For general graphs we showed there is no extremum distribution – there are many extremal distributions that are incomparable. We were able to characterize the extreme points of the polytope of all stable income distribution when only cliques of size 2 and 3 are allowed to secede.

In [23], [33], and [40] we considered energy aware scheduling problems where the goal is to minimize a convex combination of the total energy consumption and the deadline line related metric such as completion time, or flow time. Problems of this nature arise in situation where processors can be run at different speeds and there is significant energy savings when the processor speed is scaled. We showed how to extend the well-known α-point algorithms to scheduling problems with speed scaling.

Doctoral students supported

Zaiwen Wen
Necdet Aybat
Chen Chen
Yupeng Chen
Ka Chun Ma
Shiqian Ma
Zhiwei Qin
Bo Huang,

Research Associates supported

Katya Scheinberg

Publications: Appeared or to appear

[1] G. Iyengar and K-C. Ma. Cash Flow management: A Risk Management Approach. *North American Actuarial Journal.* **13**(3) (2009) 370-378.

[2] Z. Wen and D. Goldfarb. A linesearch multigrid method for large-scale nonlinear optimization. *SIAM J. on Optimization* **20** (2009) 1478-1503.

[3] D. Goldfarb and W. Yin. Parametric maximum flow algorithms for fast total variation minimization. *SIAM J. on Scientific Computing* **31** (2009) 3712-3743.

[4] D. Goldfarb, Z. Wen and W. Yin. A curvilinear search method for p-harmonic flows on spheres. *SIAM J. on Imaging Sciences* **2** (2009) 84-109

[5] L. Chen and D. Goldfarb. An interior-point piecewise linear penalty method for nonlinear programming. *Mathematical Programming* **128** (2011) 73-122.

[6] G. Iyengar and K-C. Ma. A behavioral finance based tick-by-tick model for price and volume. *Journal of Computational Finance.* **14**(1) (2010) 5780.

[7] G. Iyengar and K-C. Ma. Robust pension fund management. Journal of Asset Management. **11** (2010) 163-177.

[8] Z. Wen, W. Yin, D. Goldfarb and Y. Zhang. A fast algorithm for sparse reconstruction based on shrinkage, subspace optimization and continuation. *SIAM J. on Scientific Computing* **32** (2010) 1832-1857.

[9] Z. Wen, D. Goldfarb and W. Yin. Alternating Direction Augmented Lagrangian Methods for Semidefinite Programming. *Mathematical Programming Computation.* **2**(210) (2010) 203-230.

[10] S. Ma, D. Goldfarb, and L. Chen. Fixed point and Bregman iterative methods for matrix rank minimization. *Mathematical Programming* **128** (2011) 321-353.

[11] N. S. Aybat and G. Iyengar . A first-order penalty method for compressed sensing. *SIAM J. Optimization.* **21** (2011) 287-313.

[12] N. S. Aybat and G. Iyengar. A first-order augmented Lagrangian method for compressed sensing. *SIAM J. Optimization.* **22** (2011) 429-459. (Runner-up, 2010 INFORMS Computing Society Student Paper Award, 2010)

[13] W. Kets, G. Iyengar, S. Bowles and R. Sethi. Extremal Inequality in Stable Networks. *Games and Economic Behavior.* **73** (2011) 215-226.

[14] G. Iyengar, D. Phillips, and C. Stein. Approximating semidefinite packing problems. *SIAM J. Optimization.* **21** (2011) 231-268. (Featured Article)

[15] D. Goldfarb and S Ma. Convergence of fixed point continuation algorithms for matrix rank minimization. *Foundations of Computational Mathematics* **11** (2011) 183-210.

[16] G. Iyengar and K-C. Ma. Fast gradient descent method for mean-CVaR optimization. *Annals of OR.* **205**(1) (2012) 203-212.

[17] Z. Qin and D. Goldfarb. Structured sparsity via alternating direction methods. *J. Machine Learning Research.* **13** (2012) 1435-1468.

[18] Z. Wen, W. Yin, H. Zhang and D. Goldfarb. On the convergence of an active-set method for ℓ_1 minimization. *Optimization Methods and Software.* **27** (2012) 1127-1146.

[19] D. Goldfarb and S. Ma. Fast multiple splitting algorithms for convex optimization. SIAM J. on Optimization. **22** (2012) 533-556

[20] D. Goldfarb, S. Ma, and K. Scheinberg. Fast alternating linearization methods for minimizing the sum of two convex functions. *Mathematical Programming.* **141** (2013) 349-382.

[21] B. Huang, S. Ma and D. Goldfarb, . Accelerated linearized Bregman methods. *J. of Scientific Computing.* **54** (2-3) (2013) 428-453.

[22] Z. Qin, K. Scheinberg, and D. Goldfarb. Efficient block-coordinate descent algorithms for the *group LASSO.* Mathematical Programming Computation. **5** (2013) 143-169.

[23] R. Carrasco, G. Iyengar and C. Stein. Single Machine Scheduling with Job-Dependent Convex Cost and Arbitrary Precedence Constraints. *Oper. Res. Letters.* **41**(5) (2013) 436-441.

[24] C. Chen, G. Iyengar and C. C. Moallemi. An axiomatic approach to systemic risk. *Management Science.* **59**(6) (2013) 373-1388. (Honorable Mention, INFORMS George Nicholson Student Paper Competition, 2011)

[25] D. Goldfarb, K. Scheinberg and X. Bai. Fast first-order methods for composite convex optimization with backtracking. *Foundations of Computational Mathematics.* **14** (2014) 389-417.

[26] S. N. Aybat, D. Goldfarb, and S. Ma. Efficient algorithms for robust and stable principal component pursuit problems. *Computational Optimization and Applications.* **58** (2014) 1-29.

[27] D. Goldfarb and Z. Qin. Robust low-rank tensor recovery: models and algorithms. *SIAM J. on Matrix Analysis and Applications.* **35** (2014) 225-253.

[28] N. S. Aybat and G. Iyengar. Unified Approach for Minimizing Composite Norms. *Mathematical Programming.* **144**(1-2) (2014) 181-226. (Winner, SIAM Student Paper Prize, 2011)

[29] M. Mu, B. Huang, J. Wright and D. Goldfarb. Square Deal: lower bounds and improved relaxations for tensor recovery. To appear *ICML 2014*.

Publications: Under Review

[30] N. S. Aybat and Garud Iyengar. An Alternating Direction Method with Increasing Penalty for Stable Principal Component Pursuit. *Computational Optimization and Applications.* 2014.

[31] N. S. Aybat and G. Iyengar. An Augmented Lagrangian Method for Conic Convex Programming. Submitted to *Mathematical Programming, Series A.* 2012.

[32] C. Abad and G. Iyengar. Portfolio selection with multiple spectral risk constraints. Submitted to *Computational Finance.* 2012.

[33] R. A. Carrasco, G. Iyengar, C. Stein. Resource Cost Aware Scheduling. Submitted to *Mathematics of Operations Research.*

[34] Y. Chen, G. Iyengar and R. Iyengar. Modeling Multimodal Consumer Heterogeneity in Conjoint Analysis - A Sparse Learning Approach. Submitted to *Journal of Marketing Research.* 2012.

[35] D. Goldfarb, Z. Qin, and S. Ma. An alternating direction method for total variation denoising. Submitted to *Optimization and Software* (2013).

[36] B. Huang, C. Mu, D. Goldfarb and J. Wright. Provable low-rank tensor recovery. Submitted to *Foundations of Computational Mathematics.* 2014.

Publications: Refereed conferences

[37] D. Goldfarb, S. Ma and Z. Wen. Solving low-rank matrix completion problems efficiently. *Proceedings of IEEE Forty-Seventh Annual Allerton Conference on Communication, Control and Computing* (2009) 1013-1020.

[38] G. Iyengar, D. Phillips, and C. Stein. Feasible and accurate algorithms for covering semidefinite programs. SWAT 2010. *Acceptance rate: 36/78 = 50%*

[39] D. Goldfarb, S. Ma, and K. Scheinberg. Sparse inverse covariance selection via alternating linearization methods. *Advances in Neural Information Processing Systems* **23** (NIPS 2010). Edited by J. Lafferty, C.K.I Williams, J. Shawe-Taylor, R.S. Zemel and A. Culotta

[40] G. Iyengar, R. Carrasco, and C. Stein. Energy aware scheduling: minimizing total energy cost and completion time by α-points and α-speeds. 10th Workshop on Models and Algorithms for Planning and Scheduling Problems. MAPSP 2011.

Chapters in collections

[41] G. Iyengar. Parametric network flows and adword auctions. OPTIMA, April 2010.

[42] G. Iyengar. Robust Optimization, *Encyclopedia of Quantitative Finance.* Wiley. 2011.

[43] Z. Wen, D. Goldfarb, and K. Scheinberg. Block coordinate descent methods for semidefinite programming. In Handbook of Semidefinite, Cone and Polynomial Optimization: Theory, Algorithms, Software and Applications, edited by M. Anjos and J. Lasserre, Springer (2011) 533-564.