EXACT PRIMAL-DUAL REGULARIZATION OF LINEAR PROGRAMS

Michael P. Friedlander (University of British Columbia)
Dominique Orban (Ecole Polytechnique de Montréal)

August 19–August 28, 2006

Abstract

In the framework of linear programming, we propose a theoretical justification for regularizing the linear systems used to compute search directions when the latter are (nearly) rank deficient. Our research program is based on the analysis of a primal-dual infeasible algorithm for linear programs (LPs) with explicit primal and dual regularization. The goal is to establish a rigorous connection between proximal-point, regularization, trust-region, and augmented Lagrangian methods. The regularization is termed *exact* to emphasize that, although the LP is perturbed, we are still able to recover a solution of the original LP, independently of the values of the regularization parameters.

1 General Overview

Consider the primal-dual pair of linear programs (LPs)

$$\begin{array}{llll} \text{(P)} & \text{minimize}_x & c^T x & \text{subject to} \quad Ax = b, \quad x \geq 0, \\ \text{(D)} & \text{minimize}_{y,z} & -b^T y & \text{subject to} \quad A^T y + z = c, \quad z \geq 0, \end{array}$$

where $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and $A \in \mathbb{R}^{m \times n}$. In primal-dual interior-point methods for linear programming, the computational kernel lies in the solution of an indefinite system of linear equations to determine search directions. At each iteration, a Karush-Kuhn-Tucker (KKT) system of the form

$$\begin{bmatrix} -D & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} d_x \\ d_y \end{bmatrix} = \begin{bmatrix} f_x \\ f_y \end{bmatrix} \tag{1}$$

must be solved; the diagonal matrix D and right-hand side (f_x, f_y) change at each iteration. These systems may be solved "as is" (using direct methods or iterative methods suitable for symmetric indefinite systems), or d_x and d_y may be solved for separately by using the normal equations

$$AD^{-1}A^T d_y = AD^{-1} f_x + f_y \quad \text{and} \quad D d_x = f_x - A^T f_y. \tag{2}$$

In either case, near rank deficiency of A, or near singularity of D (diagonal elements that are equal or very close to zero), can give rise to inefficient or unstable solutions of these linear systems. Our research program consists of exploring the theoretical and numerical properties of a primal-dual regularization that perturbs (P) and (D) in ways that favourably affect the numerical properties of the linear algebra subproblems that arise in interior-point methods for LPs. Such regularizations have been used in practical implementations for solving large-scale problems (see, e.g., [AG99,ST96]). Our goal is to provide a strong theoretical justification for their use in practical implementations and also to understand the implications of regularization on the convergence properties of primal-dual interior-point methods.

Our analysis begins with LPs because of the important properties that result from their strong structure. We note, however, that our work can be extended to more general convex quadratic programs and even to nonlinear convex programs under a regularity condition. This important research topic will be the subject of our followup project.

2 Progress

During our week at BIRS, we examined several aspects of the primal-dual regularization approach. Our first task was to understand the connection of regularization and its algorithmic implications to existing methods.

Connection to augmented Lagrangians. Dual regularization via the ℓ_2 norm turns out to be equivalent to an augmented Lagrangian approach for solving (P), which is based on *approximately* solving the sequence of subproblems

$$\begin{array}{ll} \text{minimize}_x & c^T x + \tfrac{1}{2}\sigma\|r\|_2^2 + y_k^T r \\ \text{subject to} & Ax + r = b, \quad x \geq 0. \end{array} \tag{3}$$

An appropriate barrier-formulation of (3) leads to a linear system

$$\begin{bmatrix} -D & A^T \\ A & \sigma I \end{bmatrix} \begin{bmatrix} d_x \\ d_y \end{bmatrix} = \begin{bmatrix} \bar{f}_x \\ \bar{f}_y \end{bmatrix}$$

which is a perturbation of (1) and has well-defined solutions even if A is rank deficient.

A second crucial implication of the augmented-Lagrangian connection hinges on this fact: the augmented-Lagrangian algorithm, applied to LPs, is finitely convergent. We can therefore deduce the convergence of a primal-regularized or dual-regularized version of the interior-point algorithm that we propose. In particular, this connection allows us to lay the groundwork for proving fast asymptotic convergence of our algorithm. This is a refinement of the analysis carried out by Setiono who, for a similar algorithm, establishes a linear convergence rate only.

Generic regularizations. Before arriving at BIRS, we were only considering regularizations based on the ℓ_2 norm. But the connections to the augmented Lagrangian method that unfolded during our time at BIRS suggested that we should also consider much more general regularization functions. In essence, we could leverage much of the powerful analysis of augmented Lagrangian methods applied to general convex optimization problems.

We therefore began to consider the broad class of LP regularizations

$$\text{minimize}_x \quad c^T x + \phi(x) \quad \text{subject to} \quad Ax = b,\ x \geq 0, \tag{4}$$

where ϕ was any function that satisfied

$$\phi(x) - \phi(y) = \frac{\gamma}{2}\|x - y\|^2 + O(\|x - y\|^\alpha) \quad \text{for any} \quad \alpha > 1, \tag{5}$$

for some positive constant γ. This requirement of ϕ is reminiscent of the quadratic growth condition that has recently received much attention in the context of constraint qualifications for general nonlinear optimization. We anticipate that our theoretical results will carry over to such functions. In particular, the notion of a Bregman distance—well established in convex programming—seems to possess the desired properties.

Numerical experiments. We implemented a preliminary version of our regularized primal-dual algorithm within the established open-source solver GLPK, which implements both the simplex method and Mehrotra's predictor-corrector method, the latter being a popular variant of the primal-dual interior-point method. We intend to contribute our modification to GLPK back to the optimization community.

The current experiments are based on using a primal-dual ℓ_2 regularization of the classical long-step interior-point method. We have already obtained promising numerical results confirming our theoretical results: that in our regularization scheme, the precise value of the regularization parameter is inconsequential.

For testing purposes, we have selected a test set of degenerate linear programs from the Netlib collection. Those problems have been the center of attention of a large number of algorithms for linear programming and their properties are reasonably well understood by the community. We therefore believe that they form a meaningful test set for our purposes.

3 Outcome of the Meeting

The convergence theory for primal regularization on either the primal or the dual problem is complete and a realistic large-scale implementation is nearly finalized. Our work at BIRS allowed us to shed light on a few shortcomings of an extension of this theory to the case of simultaneous primal-dual regularizations. In particular, we now believe that while global convergence holds in this case, fast local convergence does not.

We intend to explore the application of our approach, with the necessary adjustments, to more general problem classes, such as convex programs, second-order cone programs and semi-definite programs. The properties that such classes share with linear programs gives good reasons to believe that most results will still hold.

References

[AG99] A. Altman and J. Gondzio. Regularized symmetric indefinite systems in interior point methods for linear and quadratic optimization. *Optimization Methods and Software*, 11(12):275–302, 1999.

[MM79] O. L. Mangasarian and R. R. Meyer. Nonlinear perturbation of linear programs. *SIAM J. Control and Optimization*, 17(6):745–752, 1979.

[Roc73] R. T. Rockafellar. The multiplier method of Hestenes and Powell applied to convex programming. *Journal of Optimization Theory and Applications*, 12:555–562, 1973.

[Roc76a] R. T. Rockafellar. Augmented lagrangians and applications of the proximal point algorithm in convex programming. *Mathematics of Operations Research*, 1(2):97–116, 1976.

[Roc76b] R. T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM J. Control and Optimization*, 14(5):877–898, 1976.

[Roc97] R. T. Rockafellar. *Convex Analysis*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, New Jersey, 1997.

[Sau96] M. A. Saunders. Cholesky-based methods for sparse least squares: The benefits of regularization. In L. Adams and J. L. Nazareth, editors, *Linear and Nonlinear Conjugate Gradient-Related Methods*, pages 92–100. SIAM, 1996.

[Sau99] M. A. Saunders. Solution of sparse linear equations using cholesky factors of augmented systems. SOL Report 99-1, Dept. of EESOR, Stanford University, 1999.

[ST96] M. A. Saunders and J. A. Tomlin. Solving regularized linear programs using barrier methods and KKT systems. SOL Report 96-4, Dept. of EESOR, Stanford University, 1996.

[Set92] R. Setiono. Interior proximal point algorithm for linear programms. *J. Optim. Th. and Appl.*, 74(3):425–444, 1992.

[Van95] R. J. Vanderbei. Symmetric quasi-definite matrices. *SIAM J. Optimization*, 5(1):100–113, 1995.

[Wri97] S. J. Wright. *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia, USA, 1997.