Journal of the Applied Mathematics, Statistics and Informatics (JAMSI), **5** (2009), No. 1

INCONSISTENT LINEAR INEQUALITIES, GENERALIZED NEWTON METHOD AND CONIC OPTIMIZATION

M. SALAHI AND M. PARANDEGAN

Abstract

In this paper, our focus is on a C^1 convex minimization problem. First we present its dual and discuss the relation between the solutions of primal and dual problems. Then we introduce a Newton like algorithm to solve it. Moreover, a conic optimization reformulation of this convex problem is introduced. Finally, the practical efficiency of Newton like algorithm is compared to the conic optimization framework on several randomly generated problems.

Mathematics Subject Classification 2000: 15A39, 90C25
Additional Key Words and Phrases: Newton method, Conic optimization, Inconsistent linear inequalities, Convex optimization

1. INTRODUCTION

Inconsistent set of linear inequalities might frequently arise in real world problems [Censor et al. 2008]. It might be due to various reasons such as errors in data, wrong formulation, lack of communications between groups who are responsible for the modeling, and many others. To avoid beginning the modeling from scratch, or repeating the experiments due to high costs, it makes sense to correct such systems to consistent systems by minimal changes in problem data [Amaral and Barahona 2005]. Throughout this paper we deal with the following inconsistent set of linear inequalities:

$$Ax \leq b. \tag{1}$$

Furthermore, we consider correcting such systems to consistent systems with minimal changes just in the right hand side vector b. However, considering both A and b is more general and has been studied in [Salahi and Ketabchi 2009]. Thus doing the minimal correction using the right hand side vector in l_2 norm is equivalent to solve the following minimization problem:

$$\min_{x \in R^n} ||(Ax - b)_+||^2. \tag{2}$$

The objective function in (2) is C^1 [Hiriat-Urruty et al. 1984] and let us denote it by $f(x)$. Its gradient is given by $\nabla f(x) = 2A^T(Ax - b)_+$. Although the hessian is not defined for this function, but the generalized Hessian in the sense of Clarke [Clarke 1976], [Hiriat-Urruty et al. 1984] is defined and has many properties of the classical Hessian. The generalized Hessian is denoted by $\partial^2 f(x)$, and is the set of all matrices M of the form

$$M = 2A^T DA, \tag{3}$$

where D is a diagonal matrix with D_{ii} equal to one when $(Ax - b)_i > 0$, it is zero when $(Ax - b)_i < 0$, and D_{ii} is in $[0, 1]$ when $(Ax - b)_i = 0$. Obviously, for any choice of diagonal matrix D, the generalized Hessian is positive semidefinite, thus $f(x)$ is a piecewise convex function. In this paper, in our implementation we let $D_{ii} = 0$ when $(Ax - b)_i = 0$ and call it H.

The dual of (2) is given by

$$\begin{aligned} \min \quad & 2b^T u + \|u\|^2 \\ & A^T u = 0, \\ & u \geq 0, \end{aligned} \tag{4}$$

which is a strictly convex minimization problem. The following theorem gives the relation between primal and dual solutions.

Lemma 1.1. *Let x^* be an optimal solution for problem (2). Then $u^* = (Ax^* - b)_+$ is an optimal solution for (4).*

Proof. Since x^* is optimal for (2) and it is a convex minimization problem, thus $A^T(Ax^* - b)_+ = 0$ and $(Ax^* - b)_+ \geq 0$. Furthermore, the objective functions of primal and dual are equal at x^* and $u^* = (Ax^* - b)_+$, thus u^* is optimal for (4). □

2. GENERALIZED NEWTON METHOD

In this section we present a Newton like algorithm for solving (2) and call it generalized Newton method [Salahi and Ketabchi 2009]. Analogous to the Newton method, the following direction is chosen and a step along it with an appropriate step size is taken:

$$d := (H + \delta I)^{-1} \nabla f(x),$$

where δ is a small positive constant like $\delta = 10^{-4}$. Since H is positive semidefinite, then this slight perturbation of generalized Hessian makes it invertible and thus the definition of d is well defined.

Generalized Newton Algorithm

Inputs: An accuracy parameter $\epsilon > 0$;
A perturbation parameter $\delta = 10^{-4}$;
A starting point $x_0 \in R^n$;
begin
$i = 0$;
while $\|\nabla f(x_i)\| \geq \epsilon$
$x_{i+1} = x_i - \alpha_i (H_i + \delta I)^{-1} \nabla f(x_i)$,
where α_i is the Armijo step size satisfying
$\alpha_i = \max\{1, \frac{1}{2}, \frac{1}{4}, \cdots\}$ for which
$f(x_{i+1}) - f(x_i) \geq -\frac{\alpha_i}{4} \nabla f(x_i)^T d_i$;
$i = i + 1$;
end
end

Remark 2.1. For any iterate of the Generalized Newton Algorithm we have

$$-\nabla f(x_i)^T d_i = \nabla f(x_i)^T (H_i + \delta I)^{-1} \nabla f(x_i) \geq (\delta + \left\|A^T A\right\|)^{-1} \left\|\nabla f(x_i)\right\|^2$$

Using this inequality the global convergence follows from Theorem 2.1 in [Mangasarian 1995].

3. CONIC OPTIMIZATION REFORMULATION

Recently optimization over symmetric cones, such as second order cones and cone of positive semidefinite matrices are widely used and many real world problems are modeled using these cones. In this section we give an alternative formulation of (2) using second order and linear cones. Before that let us first briefly introduce second order cones and optimization over these cones, called second order conic programming (SOCP) [Alizadeh and Goldfarb 2003].

Definition 3.1. A second order (Lorentz) cone is denoted by Q_n and is defined as

$$Q_n = \{(x_1, x_2, \cdots, x_n) | \sqrt{x_2^2 + \cdots, x_n^2} \leq x_1\}.$$

Analogous to the nonnegative orthant as a cone, Q_n posses the following fundamental properties. These properties are crucial in developing interior point algorithms for solving conic optimization problems.

—Q_n is a closed and convex cone.
—Q_n is self dual.
—Q_n is pointed and has nonempty interior.

Let us denote by K the product of nonnegative and second order cones. A primal standard form for a conic optimization problem is given by

$$\begin{aligned} \min \quad & c^T x \\ & Ax = b \\ & x \in K \end{aligned} \tag{5}$$

and its dual is given by

$$\begin{aligned} \max \quad & b^T y \\ & c - A^T y \in K. \end{aligned} \tag{6}$$

As we see, if we simply let $K = R_+^n$, then we have a linear programming with its dual. However, in conic optimization framework, K is a product of several linear and second order cones with different sizes. In other words, some part of the vector x might belong to linear and some others to second order cones, and we might have several linear and second order cones here.

LEMMA 3.2. *One has*

$$\|x_+\| \leq t \Leftrightarrow \exists y_1, \cdots, y_n : y_i \geq 0, \ x_i \leq y_i, \ \forall \ i = 1, \cdots, n, \|y\| \leq t.$$

PROOF. Suppose that $\|x_+\| \leq t$. If $t = 0$, then vector x does not have any positive components i.e., $x_i \leq 0, \ \forall i = 1, \cdots, n$. Then the right hand side inequalities hold for $y_i = 0,\ \forall i = 1, \cdots, n$. Now let $t \neq 0$, then to have right hand side true for some y, it is sufficient to set $y_i = (x_i)_+,\ \forall i = 1, \cdots, n$. Vice versa, suppose that right hand side holds. Then since $y_i \geq 0\ \forall i = 1, \cdots, n$ and $x_i \leq y_i$, thus $\|x_+\| \leq \|y\| \leq t$, which completes the proof. □

Using Lemma 3.2, (2) is equivalent to the following SOCP:

$$\begin{aligned} \min \quad & t \\ & Ax - b \leq y, \\ & ||y|| \leq t, \\ & y_i \geq 0, \forall i = 1, \cdots, m. \end{aligned} \tag{7}$$

We further can write it as the standard dual form (6):

$$\begin{aligned} \max \quad & -t \\ & \begin{bmatrix} -A & I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + b \in R_+^m, \\ & y \in R_+^m, \\ & (t, y^T) \in Q_{m+1}, \end{aligned} \tag{8}$$

which K is the product of two linear and one second order cones.

Although up to now Simplex like algorithms have not been proposed to solve conic optimization problems, but Interior Point algorithms for linear optimization have been successfully extended to handle conic optimization problems. Several high performance software packages like MOSEK, SeDumi, and SDPT3 [Andersen and Andersen 2000], [Sturm 1999], [Toh et al. 1999] have been designed and are widely used by many people from various fields of science and engineering. In the next section we use SeDumi to solve our randomly generated problems in conic form (8).

4. COMPUTATIONAL EXPERIMENTS

In this section, we compare the Generalized Newton Algorithm with SeDumi for solving (2) and (8) respectively, on several randomly generated inconsistent linear inequalities. Our experiments show that both algorithms result to the same objective value but not necessarily to the same solution norms. However the time differences specially on large problems are very significant, the Generalized Newton algorithm is extremely faster than SeDumi. Thus it can be cast as very simple classical like algorithm for solving problem (2).

REFERENCES

Alizadeh, F., and Goldfarb, D. 2003. Second order cone programming, *Math. Prog. Ser. B*, 95, 3–51.

Amaral, P., and Barahona, P. 2005. A framework for optimal correction of inconsistent linear constraint, *Constraints,* 10(1), 67–86.

Andersen, E.D., and Andersen, K.D. 2000. The MOSEK interior point optimizer for linear programming: an implementation of the homogeneous algorithm, in: H. Frenk, K. Roos, T. Terlaky, and S. Zhang, editors, *High Performance Optimization*, Kluwer Academic Publishers, 197–232.

Table I. **Comparison of Generalized Newton method and SeDumi**

m, n		$\|(Ax^* - b)_+\|$	$\|x^*\|$	$\|\nabla f(x^*)\|_\infty$	time(sec)
$50, 10$	GNewton	11.593	2.2954×10^{-1}	1.5632×10^{-13}	0.06
	SeDumi	11.593	2.2954×10^{-1}	1.4424×10^{-4}	1.1
$100, 50$	GNewton	2.7742×10^{1}	3.4484	1.3625×10^{-11}	0.08
	SeDumi	2.7742×10^{1}	3.4484	7.0432×10^{-6}	1.4
$200, 100$	GNewton	3.9894×10^{1}	5.4985	2.6915×10^{-11}	0.6
	SeDumi	3.9894×10^{1}	5.4985	6.4036×10^{-6}	3
$300, 100$	GNewton	4.6506×10^{1}	6.6898×10^{-1}	7.4181×10^{-12}	0.33
	SeDumi	4.6506×10^{1}	6.6898×10^{-1}	5.3636×10^{-4}	7.5
$500, 300$	GNewton	6.4895×10^{1}	7.1003	1.2417×10^{-10}	2.2
	SeDumi	6.4895×10^{1}	1.6351×10^{1}	5.4482×10^{-6}	19.5
$600, 500$	GNewton	6.7955×10^{1}	1.4312×10^{1}	4.0059×10^{-10}	6.7
	SeDumi	6.7955×10^{1}	3.2997×10^{1}	1.3796×10^{-4}	68.5
$1000, 300$	GNewton	8.6804×10^{1}	1.0559	7.7901×10^{-10}	2.1
	SeDumi	8.6804×10^{1}	1.0559	1.9297×10^{-3}	202

Censor, Y., Ben-Israel, A., Xiao, Y., and Galvin, J.M. 2008. On linear infeasibility arsing in intensity-modulated radiation teraphy inverse planning, *Linear Algebra and Its Applications,* 428, 1406–1420.

Clarke, F.H. 1976. On the inverse function theorem. *Pacif J Math,* 64, 97–102.

Hiriat-Urruty,J.B., Strodiot, J.J., and Nguyen, V.H. 1984. Generalized Hessian marices and second order optimality conditions for problems c^{L1} data, *Applied Mathematics and Optimization*, 11, 43–56.

Mangasarian, O. 1995. Parallel gradient distribution in unconstrained optimization, *SIAM Journal On Optimization,* 33(6), 1916–1925.

Salahi, M., and Ketabchi, S. 2009. Correcting an inconsistent set of linear inequalities by generalized Newton method, accepted for publication in *Optimization Methods and Software.*

Sturm, J.F. 1999. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones, *Optimization Methods and Software*, 11/12, pp. 625–653. Special issue on Interior Point Methods (CD supplement with software).

Toh, K.C., Todd, M.J., and Tutuncu, R.H. 1999. SDPT3– a Matlab software package for semidefinite programming, *Opti. Methods and Soft.*, 11, 545–581.

M. SALAHI, M. PARANDEGAN
Department of Mathematics,
University of Guilan,
Rasht, Iran,
e-mail: salahim@guilan.ac.ir, salahi.maziar@gmail.com.

Received February 2009