AMO - Advanced Modeling and Optimization, Volume 16, Number 2, 2014

Nonnegative least squares solution in inequality sense via LCP reformulation

Maziar Salahi* and Akram Taati*

Abstract

In this paper, a monotone linear complementarity reformulation of the nonnegative least squares problems in inequality sense is introduced. Then Mehrotra's predictor-corrector interior point algorithm is applied to solve the resulting monotone linear complementarity problem. Our numerical experiments on several randomly generated test problems show that Mehrotra's algorithm overperforms the widely used generalized Newton-penalty method.

Keywords: Linear inequalities, Least squares problems, Monotone linear complementarity problem, Interior-point methods, Mehrotra's predictor-corrector algorithm.

AMS subject classification 2010: 90C25; 90C51; 49M15

1. Introduction

Mathematical models described by linear systems arise in many areas such as radiation treatment planning, statistical analysis, training neural networks and classification . In practice, the resulting systems are often inconsistent. Several algorithms are proposed in order to make them feasible. A well studied approach for correcting an inconsistent system to a consistent one, is to make the least changes in problem data. For example, in [1] the authors obtain an optimal correction of an inconsistent linear system, where only the nonzero coefficient of the constraint matrix are allowed to be perturbed for reconstructing a consistent system. In [9] the authors have applied the changes simultaneously in the coefficient matrix and the right hand side vector and have provided a feasible

*Department of Applied Mathematics, Faculty of Mathematical Sciences, University of Guilan, Rasht, Iran, salahim@guilan.ac.ir

AMO - Advanced Modeling and Optimization. ISSN: 1841-4311

direction method for solving the fractional nonconvex minimization problem results from the optimum correction of the inconsistent linear system. In another study, the authors have considered the changes in the coefficient matrix only and for solving the resulting fractional quadratic problem, they have utilized Genetic algorithm [10]. The optimal correction of an inconsistent set of linear inequlities using the l_2 norm by minimal changes in the right hand side vector is also well studied in [11,14,16] where the authors have applied generalized Newton method, and interior point gradient algorithm to solve the underlying convex problem.

In this paper, we consider the following inconsistent set of linear inequalities:

$$\begin{aligned} Ax &\leq b, \\ x &\geq 0, \end{aligned}$$

where $A \in \mathbb{R}^{m\times n}$, with $m \geq n$ and $b \in \mathbb{R}^m$. Correcting such system to a consistent system with minimal changes in the right hand side vectors in l_2 norm is equivalent to solve the following minimization problem [11]:

$$\min_{x\geq 0} \ \frac{1}{2}\|(Ax-b)_+\|^2 \qquad (1)$$

where $a_+ = \max(a, 0)$. Problem (1) is a convex minimization problem and its KKT optimality conditions are as follow:

$$\begin{aligned} &A^T(Ax-b)_+ - y = 0, \\ &x^T y = 0, \qquad (2) \\ &x \geq 0, \quad y \geq 0, \end{aligned}$$

where y is the Lagrange multiplier corresponding to the inequality constraint.

It is well known that modern primal-dual interior point methods provide very efficient solution techniques for linear and nonlinear optimization problems [18]. Under the assumption of smoothness on objective function and constraints, primal-dual interior point algorithms are extended to general nonlinear programming [5,17]. However, the objective function of (1) is nonsmooth [6], thus the existing primal-dual interior point methods can not be directly applied to solve it. In this paper, we show that KKT optimality conditions (2) is equivalent to a monotone linear complementarity problem (LCP). Then we apply Mehrotra's predictor –corrector interior point algorithm to solve the resulting monotone LCP. We have compared Mehrotra's predictor-corrector algorithm with the so called Generalize Newton penalty algorithm on various randomly generated test problems. Our numerical experiments show that Mehrotra's predictor-corrector algorithm finds an optimal solution much faster than the Generalized Newton algorithm.

Notations

For a given vector $x \in \mathbb{R}^n$, $X = diag(x)$ is a diagonal matrix with x_i's as diagonal elements and e denotes the vector of all ones.

2. LCP Reformulation

The LCP is to find a pair $z, w \in \mathbb{R}^n$ such that

$$\begin{aligned} Mz + q &= w, \\ z^T w &= 0, \qquad \text{LCP(M,q)} \\ z \geq 0, w &\geq 0, \end{aligned}$$

where $q \in \mathbb{R}^n$ and $M \in \mathbb{R}^{n\times n}$ is a positive semidefinite matrix.

Theorem 1.The optimality conditions (2) is equivalent to a monotone LCP.

Proof. Let the pair (x, y) satisfies KKT conditions (2) . By introducing $z = (Ax - b)_+$, we obtain the following system:

$$\begin{aligned} &A^T z - y = 0, \\ &2z = |Ax - b| + (Ax - b), \\ &x^T y = 0, \\ &x \geq 0, y \geq 0. \end{aligned}$$

Now we split $(Ax - b)$ to its positive part $z_1 = \max(Ax - b, 0)$ and its negative part $z_2 = \max(-Ax + b, 0)$. Thus $(Ax - b) = z_1 - z_2$ and $|Ax - b| = z_1 + z_2$ with $z_1, z_2 \geq 0$ and ${z_1}^T z_2 = 0$. Therefore (2) is equivalent to the following system

$$\begin{aligned} &A^T z_1 - y = 0, \\ &Ax - b = z_1 - z_2 \\ &x^T y = 0, \ \ {z_1}^T z_2 = 0, \\ &x \geq 0, y \geq 0, z_1 \geq 0, \ z_2 \geq 0, \end{aligned} \tag{3}$$

which is a monotone LCP(M, q) with

$$M = \begin{bmatrix} 0_{n\times n} & A^T \\ -A & I \end{bmatrix},\ q = \begin{bmatrix} 0_{n\times 1} \\ b \end{bmatrix}, z = \begin{bmatrix} x \\ z_1 \end{bmatrix},\ w = \begin{bmatrix} y \\ z_2 \end{bmatrix}.$$

In the next section, we discuss Mehrotra's predictor-corrector interior point algorithm to solve the monotone LCP.

3 . Mehrotra's Predictor-Corrector Algorithm

There are several approaches to solve LCP's which have been studied intensively [2,7,12,13,18]. Among them, interior-point methods (IPMs) are one of the most remarkable methods [3,7,8,18,19]. In this section, we apply the infeasible version of Mehrotra's predictor-corrector (MPC) algorithm for solving the LCP reformulation (3) as follow:

$$\begin{bmatrix} 0_{n\times n} & A^T \\ -A & I \end{bmatrix} \begin{bmatrix} x \\ z_1 \end{bmatrix} + \begin{bmatrix} 0_{n\times 1} \\ b \end{bmatrix} = \begin{bmatrix} y \\ z_2 \end{bmatrix},$$

$$z := \begin{bmatrix} x \\ z_1 \end{bmatrix} \geq 0, w := \begin{bmatrix} y \\ z_2 \end{bmatrix} \geq 0,\ \ z^T w = 0. \quad (4)$$

Based on Mehrotra's predictor –corrector strategy, first a predictor direction $(\Delta z^{aff}, \Delta w^{aff})$ is obtained by solving the following system which corresponds computing the pure Newton direction for the system (4) :

$$\begin{bmatrix} 0 & -I & A^T & 0 \\ -A & 0 & I & -I \\ 0 & 0 & Z_2 & Z_1 \\ Y & X & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x^{aff} \\ \Delta y^{aff} \\ \Delta {z_1}^{aff} \\ \Delta {z_2}^{aff} \end{bmatrix} = \begin{bmatrix} -r_1 \\ -r_2 \\ -Z_1 Z_2 e \\ -XYe \end{bmatrix} \quad (5)$$

where $r_1 = A^T z_1 - y$, $r_2 = Ax - b - z_1 + z_2$ are the residual and $X = diag(x), Y = diag(y), Z_1 = diag(z_1), Z_2 = diag(z_2)$. This direction is often called the affine –scaling direction. It is easily seen that the coefficient matrix in (5) is nonsingular whenever $z > 0$ and , $w > 0$. It should be noted the solution of (5) is obtained as follow rather than solving it directly:

$$
\begin{aligned}
(A^T HA + X^{-1}Y)\Delta x^{aff} &= -r_1 - A^T Hr_2 + A^T Hz_2 - y,\\
\Delta y^{aff} &= -y - X^{-1}Y\Delta x^{aff},\\
\Delta {z_1}^{aff} &= H(r_2 - z_2 + A\Delta x^{aff}),\\
\Delta {z_2}^{aff} &= -z_2 - {Z_1}^{-1} Z_2 \Delta {z_1}^{aff}.
\end{aligned}
$$

where $H = (I + {Z_1}^{-1} Z_2)^{-1}$. Once the affine –scaling direction is computed, the maximum step size in this direction is computed i.e., the largest α_p^{aff} and α_d^{aff} such that

$$(z + \alpha_p^{aff}\Delta z^{aff}, w + \alpha_d^{aff}\Delta w^{aff}) \geq 0,$$

where $\Delta z^{aff} \coloneqq \left(\Delta x^{aff}, \Delta {z_1}^{aff}\right)$ and $\Delta w^{aff} \coloneqq \left(\Delta y^{aff}, \Delta {z_2}^{aff}\right)$. Using the information from the predictor step, MPC algorithm computes the corrector direction by solving the following system:

$$
\begin{bmatrix} 0 & -I & A^T & 0 \\ -A & 0 & I & -I \\ 0 & 0 & Z_2 & Z_1 \\ Y & X & 0 & 0 \end{bmatrix}
\begin{bmatrix} \Delta x^{cor} \\ \Delta y^{cor} \\ \Delta {z_1}^{cor} \\ \Delta {z_2}^{cor} \end{bmatrix}
=
\begin{bmatrix} 0 \\ 0 \\ \sigma\mu e - \Delta {Z_1}^{aff} \Delta {Z_2}^{aff} e \\ \sigma\mu e - \Delta X^{aff} \Delta Y^{aff} e \end{bmatrix} \quad (6)
$$

where $\Delta {Z_1}^{aff} = diag(\Delta {z_1}^{aff})$, $\Delta {Z_2}^{aff} = diag(\Delta {z_2}^{aff})$, $\Delta X^{aff} = diag(\Delta x^{aff})$, $\Delta Y^{aff} = diag(\Delta y^{aff})$ and the centering parameter σ is chosen according to the following heuristic due to Mehrotra

$$\sigma = (\frac{\mu^{aff}}{\mu})^3,$$

where $\mu = \frac{z^T w}{(n+m)}$ and $\mu^{aff} = \frac{(z+\alpha_p^{aff}\Delta z^{aff})^T (w+\alpha_d^{aff}\Delta w^{aff})}{(n+m)}$.

Once the predictor and corrector steps are computed, they are added together to produce the composite predictor-corrector direction

$$(\Delta z, \Delta w) = \left(\Delta z^{aff}, \Delta w^{aff}\right) + (\Delta z^{corr}, \Delta w^{corr}).$$

Finally, the direction $(\Delta z, \Delta w)$ is used to generate the new iteration of the algorithm. The detailed algorithm is outlined as follow:

MPC Algorithm

Choose $(z^0, w^0) > 0$ and set $k = 0$.

While $\left\|x^k \circ \nabla q(x^k)\ \right\|_\infty > \varepsilon$ or $\min \left(\nabla q(x^k)\right) < -\varepsilon$ **do**

Step1. Set $(z, w) = (z^k, w^k)$ and solve system (5) for $(\Delta z^{aff}, \Delta w^{aff})$. Compute

$$\alpha_p^{aff} = \text{argmax}\{\alpha \in [0,1] \big| z + \alpha \Delta z^{aff} \geq 0\},$$

$$\alpha_d^{aff} = \text{argmax}\{\alpha \in [0,1] \big| w + \alpha \Delta w^{aff} \geq 0\},$$

$$\mu^{aff} = \frac{(z+\alpha_p^{aff} \Delta z^{aff})^T (w+\alpha_d^{aff} \Delta w^{aff})}{(n+m)}.$$

Step 2. Set $\sigma = (\frac{\mu^{aff}}{\mu})^3$ and solve system (6) for $(\Delta z^{corr}, \Delta w^{corr})$.

Step 3. Set $(\Delta z, \Delta w) = \left(\Delta z^{aff}, \Delta w^{aff}\right) + (\Delta z^{corr}, \Delta w^{corr})$ and compute

$$\alpha_p = argmax\{\alpha \in [0,1] | z + \alpha \Delta z > 0\},$$

$$\alpha_d = argmax\{\alpha \in [0,1] | w + \alpha \Delta w > 0\},$$

Step 5. Set $z^{k+1} = z^k + \alpha_p \Delta z$, $w^{k+1} = w^k + \alpha_d \Delta w^k$, $k = k + 1$ and go to Step 1.

End

4. Numerical Results

In this section, we compare MPC algorithm with the so-called generalized Newton-penalty (GNP) algorithm of [14,15] on several randomly generated test problems. Test problems are generated using the following MATLAB code:

```
% Generates random inconsistent system Ax<=b.
% Input: m,n,d(density)  Output: A,b.
pl=inline('(abs(x)+x)/2');
m1=max(m-round(0.5*m),m-n);
A1=sprand(m1,n,den);A1=(A1-0.5*spones(A1));
x=spdiags(rand(n,1),0,n,n)*(rand(n,1)-rand(n,1));
```

```
x=spdiags(ones(n,1)-sign(x),0,n,n)*10*(rand(n,1)-
rand(n,1));
m2=m-m1;u=randperm(m2);A2=A1(u,:);
b1=A1*x+spdiags((rand(m1,1)),0,m1,m1)*ones(m1,1);
b2=b1(u)+spdiags((rand(m2,1)),0,m2,m2)*10*ones(m2,1);
A=100*[A1;-A2]; b=[b1;-b2];
```

In our testing, ε and maximum number of iterations in both algorithms are set to 10^{-6} and 100, respectively. The penalty parameter M in GNP algorithm is considered

$$M^k = \min\left(10^{k+3}, 10^{10}\right),$$

where k is the iteration number. The average of 10 runs for both algorithms on generated examples are reported in Tables 1 and 2. One can observe that MPC algorithm finds an optimal solution extremely faster than the GNP Algorithm, specially for large-scale problems.

Table1: Comparison of MPC and GNP algorithms on several test problems with $density = 0.1$ and.

m, n		$\|x^* \circ \nabla q(x^*)\|_\infty$	$\min\ (\nabla q(x^*))$	Time (sec)	Iter
$700, 500$	MPC	3.3799×10^{-8}	-1.6458×10^{-7}	0.692	17.7
	GNP	1.4409×10^{-3}	-7.3596×10^{-3}	3.981	100
1000,500	MPC	1.1715×10^{-8}	-3.5062×10^{-7}	0.773	17.0
	GNP	2.3568×10^{-3}	-1.4575×10^{-2}	4.370	100
1500,1000	MPC	3.2447×10^{-8}	-1.8908×10^{-7}	3.948	19.3
	GNP	6.9096×10^{-3}	-3.176×10^{-2}	20.33	100
2000,1000	MPC	1.1037×10^{-8}	-6.6336×10^{-8}	4.604	19.3
	GNP	8.4747×10^{-3}	-5.1554×10^{-2}	22.51	100
2500,2000	MPC	2.6733×10^{-8}	-1.2049×10^{-7}	22.85	19.1
	GNP	2.0459×10^{-2}	-7.4667×10^{-2}	110.7	100
3000,2000	MPC	3.3383×10^{-8}	-1.5770×10^{-7}	22.87	19.4
	GNP	2.2558×10^{-2}	-1.0369×10^{-1}	109.0	100
4000,3000	MPC	3.0006×10^{-8}	-1.2703×10^{-7}	63.46	19.5
	GNP	4.7299×10^{-2}	-1.9247×10^{-1}	292.9	100

Table2: Comparison of MPC and GNP algorithms on several test problems with $density = 0.01$ and.

m, n		$\|x^* \circ \nabla q(x^*)\|_\infty$	$\min(\nabla q(x^*))$	Time (sec)	Iter
700, 500	MPC	1.9989×10^{-6}	-1.9162×10^{-7}	1.323	52.6
	GNP	1.3562×10^{-3}	-2.2919×10^{-4}	1.580	100
1000,500	MPC	7.0082×10^{-9}	-9.6664×10^{-8}	0.269	17.8
	GNP	8.6257×10^{-5}	-3.9547×10^{-4}	1.587	100
1500,1000	MPC	4.0806×10^{-8}	-1.8268×10^{-7}	1.431	19.3
	GNP	1.7387×10^{-4}	-6.8416×10^{-4}	7.983	100
2000,1000	MPC	1.2891×10^{-8}	-1.0301×10^{-7}	1.442	18.8
	GNP	2.0124×10^{-4}	-1.2329×10^{-3}	8.221	100
2500,2000	MPC	1.7964×10^{-8}	-8.5615×10^{-8}	8.855	20.1
	GNP	4.2855×10^{-4}	-1.6269×10^{-3}	46.58	100
3000,2000	MPC	1.4490×10^{-8}	-1.0529×10^{-7}	8.364	19.4
	GNP	6.3236×10^{-4}	-2.3558×10^{-3}	45.59	100
4000,3000	MPC	3.6242×10^{-8}	-1.7865×10^{-7}	23.61	19.4
	GNP	6.8655×10^{-4}	-3.6835×10^{-3}	128.3	100

5. Conclusions

In this paper, we have studied the nonnegative least squares problems in the inequality sense. We have shown that its KKT optimality conditions is equivalent to a monotone linear complenetarity problem which can be efficiently solved using Mehrotra's predictor–corrector interior point algorithm. Finally, the performance of the so called Generalized Newton-penalty algorithm and Mehrotra's predictor-corrector algorithm are compared on several randomly generated test problems. Our computational comparison show that Mehrotra's algorithm is extremely faster.

References

[1] P. Amaral, L. M. Fernandes, J. Júdice, and H. D. Sherali, On optimal zero-preserving corrections for inconsistent linear systems, Journal of Global Optimization, 45, 645-666, 2009.

[2] B. H. Ahn, Solution of nonsymmetric linear complementarity problems by iterative methods, Journal of Optimization Theory and Applications , 33(2), 175-185, 1981

[3] J. F. Bonnans and C. C. Gonzaga, Convergence of interior point algorithms for the monotone linear complementarity problem, Mathematics *of Operations Research*,21(1), 1-15, 1996.

[4] J.W. Chinneck, Feasibility and Infeasibility in Optimization: Algorithms and Computational Methods, Springer, 2008.

[5] A. El-Bakry, R. Tapia, T. Tsuchiya and Y. Zhang, On the formulation and theory of the Newton interior-point method for nonlinear programming, Journal of Optimization Theory and Applications, 89, 507-541, 1996.

[6] F. Facchinei, Minimization of SC^1 functions and the Maratos effect, Operations Research Letters, 17, 131-137, 1995.

[7] C. Gonzaga, The largest step path following algorithm for monotone linear complementarity problems, Mathematical Programming , 76(2), 309-332, 1997.

[8] Z.H. Huang, Polynomiality of high-order feasible interior point method for solving the horizontal linear complementarity problems, Journal of System Science and Mathematical Sciences , 20(4), 432-438, 2000.

[9] S. Ketabchi, H. Moosaei, Optimal error correction and methods of feasible directions, Journal of Optimization Theory and Application, 154, 209–216, 2012.

[10] S. Ketabchi, H. Moosaei, S. Fallahi, Optimal correction of infeasible system in linear equality via Genetic algorithm, Applications and Applied Mathematics: An International Journal, 5(2),1585-1591, 2010.

[11] S. Ketabchi and M. Salahi, Correcting inconsistency in linear inequalities by minimal change in the right hand side vector, Computer Science Journal of Moldova, 17(2), 179-192, 2009.

[12] C. E. Lemke, Bimatrix equilibrium points and mathematical programming, management science1,681-689, 1965

[13] O. L. Mangasarian, solution of symmetric linear complementarity problems by iterative methods, Journal of Optimization Theory and Applications , 22,465-485,1977.

[14] M. Salahi and S. Ketabchi, Correcting an inconsistent set of linear inequalities by generalized Newton method, Optimization Methods and Software, 25(3),457-465, 2010.

[15] M. Salahi and M. Ganji, A generalized Newton-penalty algorithm for large scale ill-conditioned quadratic problems, Applications and Applied Mathematics, 4(2), 273-278, 2009.

[16] M. Salahi, A. Taati, Interior point gradient algorithm for totally nonnegative least squares problems in inequality sense, to appear in International Journal of Computer Mathematcs, 2014.

[17] R. Vanderbaei and D. F. Shanno, An interior-point algorithm for nonconvex nonlinear programming, Computational Optimization and Applications, 13, 231-252, 1999.

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

[19] L. Zhang, Y. Xu, A full-Newton step primal-dual Interior point algorithm for linear complementarity problems, Journal of Information and Computing Science, 5(4), 305-313, 2010.