An Interior Point Potential Reduction Algorithm for the Linear Complementarity Problem

Masakazu Kojima* Nimrod Megiddo† Yinyu Ye‡

(Revised, December 1989)

Keywords: linear complementarity, P-matrix, interior-point, potential function, linear programming, quadratic programming

Abbreviated title: Interior-point algorithm for the LCP

Abstract. The linear complementarity problem (LCP) can be viewed as the problem of minimizing $\boldsymbol{x}^T\boldsymbol{y}$ subject to $\boldsymbol{y} = \boldsymbol{M}\boldsymbol{x} + \boldsymbol{q}$ and $\boldsymbol{x}, \boldsymbol{y} \geq \boldsymbol{0}$. We are interested in finding a point with $\boldsymbol{x}^T\boldsymbol{y} < \epsilon$ for a given $\epsilon > 0$. The algorithm proceeds by iteratively reducing the potential function

$$f(\boldsymbol{x}, \boldsymbol{y}) = \rho \ln \boldsymbol{x}^T\boldsymbol{y} - \sum \ln x_j y_j \ ,$$

where, for example, $\rho = 2n$. The direction of movement in the original space can be viewed as follows. First, apply a *linear* scaling transformation to make the coordinates of the current point all equal to 1. Take a gradient step in the transformed space using the gradient of the transformed potential function, where the step size is either predetermined by the algorithm or decided by line search to minimize the value of the potential. Finally, map the point back to the original space.

A bound on the worst-case performance of the algorithm depends on the parameter $\lambda^* = \lambda^*(\boldsymbol{M}, \epsilon)$, which is defined as the minimum of the smallest eigenvalue of a matrix of the form

$$(\boldsymbol{I} + \boldsymbol{Y}^{-1}\boldsymbol{M}\boldsymbol{X})(\boldsymbol{I} + \boldsymbol{X}\boldsymbol{M}^T\boldsymbol{Y}^{-2}\boldsymbol{M}\boldsymbol{X})^{-1}(\boldsymbol{I} + \boldsymbol{X}\boldsymbol{M}^T\boldsymbol{Y}^{-1})$$

where $\boldsymbol{X}$ and $\boldsymbol{Y}$ vary over the nonnegative diagonal matrices such that $\boldsymbol{e}^T\boldsymbol{X}\boldsymbol{Y}\boldsymbol{e} \geq \epsilon$ and $X_{jj}Y_{jj} \leq n^2$. If $\boldsymbol{M}$ is a P-matrix, λ^* is positive and the algorithm solves the problem in polynomial time in terms of the input size, $|\log \epsilon|$, and $1/\lambda^*$. It is also shown that when $\boldsymbol{M}$ is positive semi-definite, the choice of $\rho = 2n + \sqrt{2n}$ yields a polynomial-time algorithm. This covers the convex quadratic minimization problem.

*Department of Information Sciences, Tokyo Institute of Technology, Meguro, Tokyo, Japan

†IBM Research, Almaden Research Center, San Jose, CA 95120-6099, and School of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel

‡Department of Management Sciences, University of Iowa, Iowa City, IA 52242

1. Introduction

We consider the linear complementarity problem (LCP), that is, given a rational matrix $\boldsymbol{M} \in R^{n\times n}$ and a rational vector $\boldsymbol{q} \in R^n$, find vectors $\boldsymbol{x}, \boldsymbol{y} \in R^n$ such that

$$(LCP) \qquad \begin{aligned} \boldsymbol{y} &= \boldsymbol{M}\boldsymbol{x} + \boldsymbol{q} \\ \boldsymbol{x}, \boldsymbol{y} &\geq \boldsymbol{0} \\ \boldsymbol{x}^T\boldsymbol{y} &= 0 \ . \end{aligned}$$

Equivalently, we consider the LCP as a quadratic programming problem

$$\begin{aligned} \text{Minimize} \quad & \boldsymbol{x}^T\boldsymbol{y} \\ \text{subject to} \quad & \boldsymbol{y} = \boldsymbol{M}\boldsymbol{x} + \boldsymbol{q} \\ & \boldsymbol{x}, \boldsymbol{y} \geq \boldsymbol{0} \ . \end{aligned}$$

The problem of recognizing whether the LCP has a solution is NP-complete.[1] We are interested in the problem of finding a solution in special cases where existence is guaranteed. For example, if $\boldsymbol{M}$ has positive principal minors (i.e., $\boldsymbol{M}$ is a P-matrix [1]), a unique solution exists for every $\boldsymbol{q}$. Also, the problem of computing a Nash-equilibrium point in a two-person game can be reduced to an LCP with $\boldsymbol{q} = -\boldsymbol{e} = (-1, \cdots, -1)^T$ and

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{O} & \boldsymbol{B}^T \\ \boldsymbol{A} & \boldsymbol{O} \end{bmatrix}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ have positive entries. A solution always exists for such problems. It is shown in [5] that the P-matrix LCP and the equilibrium problems cannot be NP-hard unless NP = coNP.

An ϵ-complementary solution is a pair $\boldsymbol{x}, \boldsymbol{y} \geq \boldsymbol{0}$ such that $\boldsymbol{y} = \boldsymbol{M}\boldsymbol{x} + \boldsymbol{q}$ and

$$\boldsymbol{x}^T\boldsymbol{y} \leq \epsilon \ .$$

The question we are interested in here is the existence of an algorithm for finding an ϵ-complementary solution in time which is bounded by a polynomial in the input size L (i.e., the length of the binary representation of $\boldsymbol{M}$ and $\boldsymbol{q}$) and $-\log\epsilon$. It is shown in [4] that there exists ϵ, where $-\log\epsilon$ is bounded by the input size, such that an exact solution can be easily computed from an ϵ-complementary one.

The analysis of this paper is related to the one presented by Karmarkar [3] for his linear programming algorithm. Gonzaga [2] proves for linear programming essentially the same result as the one presented in Section 7. Also, Ye [6] used potential functions in a similar way.

[1]This is because the problem of recognizing existence of $(0,1)$-solutions to $\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}$ can be set as an LCP: $y_j = 1 - x_j$, $\boldsymbol{u} = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}$, $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{v} \geq \boldsymbol{0}$, $\boldsymbol{x}^T\boldsymbol{y} = \boldsymbol{v}^T\boldsymbol{u} = 0$.

2. An invariant potential function

We assume the feasible domain intersects the positive orthant, i.e., there exist $\boldsymbol{x}, \boldsymbol{y} > \boldsymbol{0}$ such that $\boldsymbol{y} = \boldsymbol{M}\boldsymbol{x} + \boldsymbol{q}$. For brevity, we say that such points are *interior.*

We define a potential function $f(\boldsymbol{x}, \boldsymbol{y})$ at interior points by

$$f(\boldsymbol{x}, \boldsymbol{y}) = \rho \ln \boldsymbol{x}^T \boldsymbol{y} - \sum_{j=1}^{n} \ln x_j y_j$$

where $\rho = \rho(n) > n$.

Lemma 2.1. *If $f(\boldsymbol{x}, \boldsymbol{y}) \leq -K$, then $\boldsymbol{x}^T \boldsymbol{y} < e^{-\frac{K}{\rho - n}}$.*

Proof: The claim follows from

$$\begin{aligned} f(\boldsymbol{x}, \boldsymbol{y}) &= \rho \ln \boldsymbol{x}^T \boldsymbol{y} - \ln \prod_{i=1}^{n} x_j y_j \\ &\geq \rho \ln \boldsymbol{x}^T \boldsymbol{y} - \ln \left(n^{-1} \sum_{i=1}^{n} x_j y_j \right)^n \\ &= (\rho - n) \ln \boldsymbol{x}^T \boldsymbol{y} + n \ln n > (\rho - n) \ln \boldsymbol{x}^T \boldsymbol{y} \ . \end{aligned}$$

∎

We use the notation $p(L)$ for any polynomial in L. Starting from a point where the potential function value is $p(L)$, if the value is reduced during each step by at least n^{-k} for some k, then it takes a polynomial number of steps (in terms of L, n, and $-\log \epsilon$) to reach an ϵ-complementary point.

We now examine the behavior of the potential function under scaling of $\boldsymbol{x}$ and $\boldsymbol{y}$. Given any feasible interior point $(\boldsymbol{x}^0, \boldsymbol{y}^0)$, let $\boldsymbol{X}$ and $\boldsymbol{Y}$ denote diagonal matrices with the coordinates of $\boldsymbol{x}^0$ and $\boldsymbol{y}^0$, respectively, in their diagonals. Define a linear transformation of space by

$$\boldsymbol{x}' = \boldsymbol{X}^{-1} \boldsymbol{x} \quad , \quad \boldsymbol{y}' = \boldsymbol{Y}^{-1} \boldsymbol{y} \ .$$

Denote

$$\boldsymbol{M}' = \boldsymbol{Y}^{-1} \boldsymbol{M} \boldsymbol{X} \quad , \quad \boldsymbol{q}' = \boldsymbol{Y}^{-1} \boldsymbol{q} \ ,$$

and

$$w_j = x_j^0 y_j^0 \ \ (j = 1, \cdots, n) \quad \text{and} \quad \boldsymbol{w} = (w_1, \cdots, w_n)^T \ .$$

Also denote

$$\boldsymbol{W} = \boldsymbol{X} \boldsymbol{Y} \ .$$

The transformed problem is

$$\begin{aligned} \text{Minimize} \quad & \boldsymbol{x}'^T \boldsymbol{W} \boldsymbol{y}' \\ \text{subject to} \quad & \boldsymbol{y}' = \boldsymbol{M}'\boldsymbol{x}' + \boldsymbol{q}' \\ & \boldsymbol{x}', \boldsymbol{y}' \geq \mathbf{0} \ . \end{aligned}$$

Note that feasible solutions of the original problems are mapped into feasible solutions of the transformed problem:

$$\boldsymbol{y}' = \boldsymbol{Y}^{-1}\boldsymbol{y} = \boldsymbol{Y}^{-1}(\boldsymbol{M}\boldsymbol{x} + \boldsymbol{q}) = \boldsymbol{M}'\boldsymbol{x}' + \boldsymbol{q}'$$

and, obviously, the point $(\boldsymbol{x}^0, \boldsymbol{y}^0)$ is mapped into $(\boldsymbol{e}, \boldsymbol{e})$. Consider the transformed potential function

$$f'(\boldsymbol{x}', \boldsymbol{y}') = \rho \ln \boldsymbol{x}'^T \boldsymbol{W} \boldsymbol{y}' - \sum_{j=1}^{n} \ln x'_j w_j y'_j \ .$$

It is easy to verify that $f'(\boldsymbol{x}', \boldsymbol{y}') = f(\boldsymbol{x}, \boldsymbol{y})$, so a reduction in one of the functions results in the same reduction in the other one.

Obviously, the matrix $\boldsymbol{M}'$ is a P-matrix if and only if $\boldsymbol{M}$ is a P-matrix. Now, consider the special case where the current point $(\boldsymbol{x}, \boldsymbol{y})$ is on the “central path”, i.e., $\boldsymbol{X}\boldsymbol{Y} = \mu \boldsymbol{I}$ for some μ. In this case, some important structure of $\boldsymbol{M}$ is preserved in the transformation to $\boldsymbol{M}'$. For example, if $\boldsymbol{M}$ arises from a linear programming problemthen $\boldsymbol{M}$ is skew-symmetric. In this case, $\boldsymbol{M}'$ is also skew-symmetric and arises from a transformed linear programming problem. Obviously, if $\boldsymbol{M}$ is symmetric then so is $\boldsymbol{M}'$. Moreover, if $\boldsymbol{M}$ is positive semi-definitethen so is $\boldsymbol{M}'$. This implies that if $\boldsymbol{M}$ arises from a convex quadratic programming problem then so does $\boldsymbol{M}'$.

3. Getting an initial point

To guarantee a polynomial running time, we have to start the algorithm at an interior point where the value of the potential function is bounded by a polynomial $p(L)$ in the size L of the input of the problem.

We note that any LCP can be reduced to an LCP whose feasible domain intersects the positive orthant, by adding a pair of complementary variables x_0, y_0:

$$(LCP') \qquad \begin{aligned} \boldsymbol{y} = \boldsymbol{M}\boldsymbol{x} + (\boldsymbol{e} - \boldsymbol{M}\boldsymbol{e} - \boldsymbol{q})x_0 + \boldsymbol{q} \\ y_0 = x_0 \\ \boldsymbol{x}, \boldsymbol{y} \geq \mathbf{0} \\ x_0, y_0 \geq 0 \\ x_j y_j = 0 \quad (j = 0, \cdots, n) \ , \end{aligned}$$

where the choice $\boldsymbol{x} = \boldsymbol{y} = \boldsymbol{e}$, $x_0 = y_0 = 1$ yields an interior feasible solution. Moreover, the value of the potential function at this point is $\rho \ln(n+1)$. Note that this reformulation preserves the P-matrix property. We may view ϵ-complementary solutions of (LCP') as approximate solutions to (LCP). It is important to note that given an ϵ-complementary solution with $\epsilon < 2^{-L}$, a simple procedure finds an exact solution (see [4]).

Another possible initialization is as follows. If the original problem has an interior point, then any polynomial-time algorithm for linear programming can be employed to find an interior point where the value of the potential function is bounded by a polynomial in the input size. This can be done as follows. Solve the following linear programming problem:

$$(P) \qquad \begin{aligned} \text{Maximize } & \xi \\ \text{subject to } & \boldsymbol{y} = \boldsymbol{M}\boldsymbol{x} + \boldsymbol{q} \\ & \boldsymbol{x}, \boldsymbol{y} \geq \xi \boldsymbol{e} \\ & \xi \leq 1 \ . \end{aligned}$$

If the problem does not have an interior feasible point, then this fact is discovered by recognizing either that (P) is infeasible or that the optimal value ξ^* of (P) is nonpositive. Assuming $\xi^* > 0$, we have $\xi^* > 2^{-p(L)}$, and also the size of this value is bounded by a polynomial in L. Moreover, at a basic optimal solution the x_j's and y_j's have polynomial sizes, and hence the value of the potential function at such a solution is bounded by a polynomial $p(L)$. Of course, this is only a worst-case estimate and the actual initial point may not require so many bits to specify.

4. The core of the algorithm

During a single iteration the algorithm takes a step in the direction corresponding to steepest descent of the potential function in the transformed space. More precisely, the algorithm transforms the current point $(\boldsymbol{x}^k, \boldsymbol{y}^k)$ into $(\boldsymbol{e}, \boldsymbol{e})$ by linear scaling, moves to a point in the direction of steepest descent, and then transforms this point back to the original space and calls it $(\boldsymbol{x}^{k+1}, \boldsymbol{y}^{k+1})$. To simplify the presentation, we assume the current point is indeed $(\boldsymbol{e}, \boldsymbol{e})$ and the potential function has the form

$$f(\boldsymbol{x}, \boldsymbol{y}) = \rho \ln \boldsymbol{x}^T \boldsymbol{W} \boldsymbol{y} - \sum_{j=1}^{n} \ln x_j w_j y_j \ .$$

We also omit the primes from $\boldsymbol{M}'$ and $\boldsymbol{q}'$.

The gradient of f is given by

$$\nabla_x f(\boldsymbol{x}, \boldsymbol{y}) = \frac{\rho}{\boldsymbol{x}^T \boldsymbol{W} \boldsymbol{y}} \boldsymbol{W} \boldsymbol{y} - \boldsymbol{X}^{-1} \boldsymbol{e}$$

$$\nabla_y f(\boldsymbol{x}, \boldsymbol{y}) = \frac{\rho}{\boldsymbol{x}^T \boldsymbol{W} \boldsymbol{y}} \boldsymbol{W} \boldsymbol{x} - \boldsymbol{Y}^{-1} \boldsymbol{e} \ ,$$

where $\boldsymbol{X}$ and $\boldsymbol{Y}$ denote diagonal matrices containing the coordinates of $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively, in their diagonals. Denote

$$\boldsymbol{g} = \frac{\rho}{\boldsymbol{e}^T \boldsymbol{w}} \boldsymbol{w} - \boldsymbol{e} \ .$$

If $\boldsymbol{x} = \boldsymbol{y} = \boldsymbol{e}$, then

$$\boldsymbol{g}_x = \nabla_x f(\boldsymbol{e}, \boldsymbol{e}) = \frac{\rho}{\boldsymbol{e}^T \boldsymbol{W} \boldsymbol{e}} \boldsymbol{W} \boldsymbol{e} - \boldsymbol{e} = \boldsymbol{g}$$
$$\boldsymbol{g}_y = \nabla_y f(\boldsymbol{e}, \boldsymbol{e}) = \frac{\rho}{\boldsymbol{e}^T \boldsymbol{W} \boldsymbol{e}} \boldsymbol{W} \boldsymbol{e} - \boldsymbol{e} = \boldsymbol{g} \ .$$

Denote by $(\boldsymbol{\Delta x}, \boldsymbol{\Delta y})$ the projection of $\nabla f(\boldsymbol{e}, \boldsymbol{e})$ on the linear space Ω defined by $\boldsymbol{\Delta y} = \boldsymbol{M} \boldsymbol{\Delta x}$. Thus, $(\boldsymbol{\Delta x}, \boldsymbol{\Delta y})$ minimizes

$$\|\boldsymbol{\Delta x} - \boldsymbol{g}\|^2 + \|\boldsymbol{\Delta y} - \boldsymbol{g}\|^2$$

subject to

$$\boldsymbol{\Delta y} = \boldsymbol{M} \boldsymbol{\Delta x} \ .$$

It follows that

$$\boldsymbol{\Delta x} = (\boldsymbol{I} + \boldsymbol{M}^T \boldsymbol{M})^{-1} (\boldsymbol{I} + \boldsymbol{M}^T) \boldsymbol{g}$$
$$\boldsymbol{\Delta y} = \boldsymbol{M} (\boldsymbol{I} + \boldsymbol{M}^T \boldsymbol{M})^{-1} (\boldsymbol{I} + \boldsymbol{M}^T) \boldsymbol{g} \ .$$

Denote $\boldsymbol{d} = ((\boldsymbol{\Delta x})^T, (\boldsymbol{\Delta y})^T)^T$ and $\boldsymbol{h} = (\boldsymbol{g}^T, \boldsymbol{g}^T)^T$.

Proposition 4.1.

$$\boldsymbol{h}^T \boldsymbol{d} = \|\boldsymbol{d}\|^2 \ .$$

Proof: The vector $\boldsymbol{d}$ is a projection of $\boldsymbol{h}$ on a linear subspace, so $\boldsymbol{d} = \boldsymbol{P} \boldsymbol{h}$ where $\boldsymbol{P}$ is a symmetric matrix such that $\boldsymbol{P}^2 = \boldsymbol{P}$. Thus,

$$\|\boldsymbol{d}\|^2 = \boldsymbol{h}^T \boldsymbol{P}^2 \boldsymbol{h} = \boldsymbol{h}^T \boldsymbol{P} \boldsymbol{h} = \boldsymbol{h}^T \boldsymbol{d} \ .$$

∎

Fact 4.2.

$$\boldsymbol{e}^T \boldsymbol{g} = \rho - n \quad \text{and hence} \quad \|\boldsymbol{g}\| \geq \frac{\rho - n}{\sqrt{n}} \ .$$

Fact 4.3.

$$\|\boldsymbol{g}\| \leq \sqrt{\rho^2 - 2\rho + n} < \rho \ .$$

Proof: By definition, since $w_j > 0$,

$$-1 < g_j < \rho - 1 \ .$$

The vector $\boldsymbol{g}$ lies in the simplex defined by these bounds, together with $\boldsymbol{e}^T\boldsymbol{g} = \rho - n$. The maximum norm of a point in this simplex is attained at a vertex, so we have

$$\|\boldsymbol{g}\| \le \|(-1, \cdots, -1, \rho - 1)^T\| = \sqrt{\rho^2 - 2\rho + n} < \rho \ .$$

∎

Corollary 4.4.

$$\max\{\|\boldsymbol{\Delta x}\|, \|\boldsymbol{\Delta y}\|\} < \rho\sqrt{2} \ .$$

Proof: Since $\boldsymbol{d}$ is the projection of $\boldsymbol{h}$ on a linear subspace,

$$\|\boldsymbol{d}\| \le \|\boldsymbol{h}\| = \|\boldsymbol{g}\|\sqrt{2} < \rho\sqrt{2} \ ,$$

which implies the claim. ∎

We now estimate the amount of reduction Δf in the value of f as we move from $\boldsymbol{x} = \boldsymbol{y} = \boldsymbol{e}$ to a point of the form $\boldsymbol{x}' = \boldsymbol{e} - t\boldsymbol{\Delta x}$, $\boldsymbol{y}' = \boldsymbol{e} - t\boldsymbol{\Delta y}$, where $t > 0$. We would like to choose t so as to achieve a reduction of at least n^{-k} for some constant k.

Proposition 4.5.

$$\ln(\boldsymbol{e}^T\boldsymbol{W}\boldsymbol{e}) - \ln\left((\boldsymbol{e} - t\boldsymbol{\Delta x})^T\boldsymbol{W}(\boldsymbol{e} - t\boldsymbol{\Delta y})\right) \ge \frac{\boldsymbol{e}^T\boldsymbol{W}(\boldsymbol{\Delta x} + \boldsymbol{\Delta y})}{\boldsymbol{e}^T\boldsymbol{W}\boldsymbol{e}}\ t - \|\boldsymbol{d}\|^2 t^2 \ .$$

Proof: First, we have

$$\begin{aligned}\ln\left((\boldsymbol{e} - t\boldsymbol{\Delta x})^T\boldsymbol{W}(\boldsymbol{e} - t\boldsymbol{\Delta y})\right) &= \ln\left(\boldsymbol{e}^T\boldsymbol{W}\boldsymbol{e} - t\boldsymbol{e}^T\boldsymbol{W}(\boldsymbol{\Delta x} + \boldsymbol{\Delta y}) + t^2(\boldsymbol{\Delta x})^T\boldsymbol{W}\boldsymbol{\Delta y}\right) \\ &< \ln \boldsymbol{e}^T\boldsymbol{W}\boldsymbol{e} - \frac{\boldsymbol{e}^T\boldsymbol{W}(\boldsymbol{\Delta x} + \boldsymbol{\Delta y}) - t(\boldsymbol{\Delta x})^T\boldsymbol{W}\boldsymbol{\Delta y}}{\boldsymbol{e}^T\boldsymbol{W}\boldsymbol{e}}\ t \ .\end{aligned}$$

However,

$$\frac{(\boldsymbol{\Delta x})^T\boldsymbol{W}\boldsymbol{\Delta y}}{\boldsymbol{e}^T\boldsymbol{W}\boldsymbol{e}} \le \|\boldsymbol{\Delta x}\| \cdot \|\boldsymbol{\Delta y}\| \le \|\boldsymbol{d}\|^2 \ .$$

∎

Proposition 4.6. *If $\boldsymbol{M}^T\boldsymbol{g} \ne -\boldsymbol{g}$, then $\boldsymbol{h}^T\boldsymbol{d} > 0$.*

Proof:

$$\boldsymbol{h}^T\boldsymbol{d} = \boldsymbol{g}^T(\boldsymbol{I}+\boldsymbol{M})(\boldsymbol{I}+\boldsymbol{M}^T\boldsymbol{M})^{-1}(\boldsymbol{I}+\boldsymbol{M}^T)\boldsymbol{g} \ .$$

Since $\boldsymbol{I}+\boldsymbol{M}^T\boldsymbol{M}$ is positive-definite, our claim follows. ∎

Denote

$$t^* = \tfrac{1}{2}\min\{(\max_i|\Delta x_i|)^{-1}, (\max_i|\Delta y_i|)^{-1}\} \ ,$$

and notice that

$$t^* \geq \frac{1}{2\|\boldsymbol{d}\|} \ .$$

Proposition 4.7. *For $t \leq t^*$,*

$$\sum_{j=1}^{n} \ln(1-t\Delta x_j)(1-t\Delta y_j) \geq -(\boldsymbol{e}^T\boldsymbol{d})\ t - 2\|\boldsymbol{d}\|^2 t^2 \ .$$

Proof: The proof follows from the inequality

$$\ln(1-d_jt) \geq -d_jt - \tfrac{1}{2}\frac{d_j^2t^2}{(1-|d_jt|)^2}$$

which holds for t such that $|d_jt| < 1$. ∎

Corollary 4.8. *For $t \leq t^*$,*

$$\Delta f \geq \|\boldsymbol{d}\|^2 t - (\rho+2)\|\boldsymbol{d}\|^2 t^2 \ .$$

Proof: The proof follows by combining the results of Propositions 4.1, 4.5, 4.7. Note that

$$\rho\frac{\boldsymbol{e}^T\boldsymbol{W}(\boldsymbol{\Delta x}+\boldsymbol{\Delta y})}{\boldsymbol{e}^T\boldsymbol{W}\boldsymbol{e}} - \boldsymbol{e}^T\boldsymbol{d} = \boldsymbol{h}^T\boldsymbol{d} = \|\boldsymbol{d}\|^2 \ .$$

∎

Lemma 4.9. *There exists a $t < t^*$ for which*[2] *$\Delta f \geq \Omega(\|\boldsymbol{d}\|^2/\rho)$. In particular, if*

$$\|\boldsymbol{d}\|^2 \geq \frac{1}{O(n^k)}$$

for some $k \geq 0$, then

$$\Delta f \geq \frac{1}{O(\rho n^k)} \ .$$

[2] The notation $\phi(x) = \Omega(\psi(x))$ means that there exists a constant c such that $\phi(x) \geq c\psi(x)$.

Proof: The maximum of $\|\boldsymbol{d}\|^2 t - (\rho+2)\|\boldsymbol{d}\|^2 t^2$ is attained at

$$t' = \frac{1}{2\rho+4} \ .$$

Let

$$t = \min\{t', t^*\} \ .$$

If

$$t' \le t^* \ ,$$

then we have

$$\Delta f \ge \frac{\|\boldsymbol{d}\|^2}{4\rho+8} \ .$$

Otherwise,

$$\Delta f \ge \|\boldsymbol{d}\|^2 t^* - (\rho+2)\|\boldsymbol{d}\|^2 (t^*)^2 > \tfrac{1}{2}\|\boldsymbol{d}\|^2 t^* > \frac{1}{4}\|\boldsymbol{d}\| > \frac{\|\boldsymbol{d}\|^2}{4\sqrt{2}\rho} \ .$$

This implies our claim. ∎

5. Suitable matrices

In view of Lemma 4.9, we need to get good lower bounds on $\|\boldsymbol{d}\|$. Recall that $\boldsymbol{d} \ne 0$ provided $\boldsymbol{M}^T\boldsymbol{g} \ne -\boldsymbol{g}$. However, we are interested in conditions on $\boldsymbol{M}$ which imply that $\|\boldsymbol{d}\|$ is bounded away from zero, hopefully by n^{-k} for some k. Moreover, since the matrix $\boldsymbol{M}$ of the preceding section is actually $\boldsymbol{M}' = \boldsymbol{Y}^{-1}\boldsymbol{M}\boldsymbol{X}$, we need to look for conditions on $\boldsymbol{M}$ which are invariant under scaling of columns and rows by positive multipliers.

Definition 5.1. For any matrix $\boldsymbol{M}$ and any diagonal matrices (with positive diagonal entries) $\boldsymbol{X}$ and $\boldsymbol{Y}$, denote

$$\boldsymbol{S} = \boldsymbol{S}(\boldsymbol{M}, \boldsymbol{Y}^{-1}, \boldsymbol{X}) = (\boldsymbol{I} + \boldsymbol{Y}^{-1}\boldsymbol{M}\boldsymbol{X})(\boldsymbol{I} + \boldsymbol{X}\boldsymbol{M}^T\boldsymbol{Y}^{-2}\boldsymbol{M}\boldsymbol{X})^{-1}(\boldsymbol{I} + \boldsymbol{X}\boldsymbol{M}^T\boldsymbol{Y}^{-1}) \ .$$

Let $\lambda(\boldsymbol{M}, \boldsymbol{Y}^{-1}, \boldsymbol{X})$ denote the smallest eigenvalue of $\boldsymbol{S}(\boldsymbol{M}, \boldsymbol{Y}^{-1}, \boldsymbol{X})$, and let

$$\lambda_0(\boldsymbol{M}) = \inf\{\lambda(\boldsymbol{M}, \boldsymbol{Y}^{-1}, \boldsymbol{X}) : X_{jj}, Y_{jj} > 0\} \ .$$

Also, let

$$\lambda_1(\boldsymbol{M}) = \inf\{\lambda(\boldsymbol{M}, \boldsymbol{Y}^{-1}, \boldsymbol{X}) : X_{jj}, Y_{jj} > 0 \text{ and } X_{jj}Y_{jj} \le n^2\} \ .$$

Obviously, $\lambda_1(\boldsymbol{M}) \ge \lambda_0(\boldsymbol{M})$. The distinction between λ_0 and λ_1 is crucial in the context of P-matrices, as we show below.

Remark 5.2. For every $\boldsymbol{M}$, $\boldsymbol{X}$ and $\boldsymbol{Y}$, $\lambda(\boldsymbol{M}, \boldsymbol{Y}^{-1}, \boldsymbol{X}) \geq 0$ and $\lambda(\boldsymbol{M}, \boldsymbol{Y}^{-1}, \boldsymbol{X}) = 0$ if and only if $\boldsymbol{I} + \boldsymbol{Y}^{-1}\boldsymbol{M}\boldsymbol{X}$ is singular. Thus $\lambda_0(\boldsymbol{M}) \geq 0$. Note that if $\boldsymbol{M}$ has nonnegative principal minors, then $\boldsymbol{I} + \boldsymbol{R}\boldsymbol{M}\boldsymbol{C}$ is nonsingular for any nonnegative diagonal matrices $\boldsymbol{R}$ and $\boldsymbol{C}$. On the other hand, if $\boldsymbol{M}$ has a negative principal minor, there exist $\boldsymbol{X}$ and $\boldsymbol{Y}$ such that $\boldsymbol{I} + \boldsymbol{Y}^{-1}\boldsymbol{M}\boldsymbol{X}$ is singular. This can be seen as follows. Suppose $S \subseteq \{1, \cdots, n\}$ is a set of indices of columns and rows corresponding to a negative minor. Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be diagonal matrices where $X_{jj} = Y_{jj}^{-1} = t$ for $j \in S$ and $X_{jj} = Y_{jj}^{-1} = \min\{t, 1/t^2\}$ otherwise. The determinant of $\boldsymbol{I} + \boldsymbol{Y}^{-1}\boldsymbol{M}\boldsymbol{X}$ tends to 1 when t tends to 0, and to $-\infty$ as t tends to infinity. It follows that for some value of t this determinant equals zero. Since $X_{jj}Y_{jj} = 1$, it follows that if $\boldsymbol{M}$ has a negative minor, then $\lambda_1(\boldsymbol{M}) = 0$.

Proposition 5.3. *If $\rho = 2n$, then throughout the execution of the algorithm,*

$$\|\boldsymbol{d}\|^2 \geq n\lambda(\boldsymbol{M}, \boldsymbol{Y}^{-1}, \boldsymbol{X}) \geq n\lambda_1(\boldsymbol{M}) \ .$$

Proof: It follows from the analysis above (see Fact 4.2) that

$$\begin{aligned}\|\boldsymbol{d}\|^2 &= \boldsymbol{d}^T\boldsymbol{h} = (\boldsymbol{\Delta x} + \boldsymbol{\Delta y})^T\boldsymbol{g} \\ &= \boldsymbol{g}^T(\boldsymbol{I} + \boldsymbol{Y}^{-1}\boldsymbol{M}\boldsymbol{X})(\boldsymbol{I} + \boldsymbol{X}\boldsymbol{M}^T\boldsymbol{Y}^{-2}\boldsymbol{M}\boldsymbol{X})^{-1}(\boldsymbol{I} + \boldsymbol{X}\boldsymbol{M}^T\boldsymbol{Y}^{-1})\boldsymbol{g} \ . \\ &\geq \lambda(\boldsymbol{M}, \boldsymbol{Y}^{-1}, \boldsymbol{X})\|\boldsymbol{g}\|^2 \geq n\lambda(\boldsymbol{M}, \boldsymbol{Y}^{-1}, \boldsymbol{X})\end{aligned}$$

To prove that actually $\lambda_1(\boldsymbol{M})$ can be used, recall that the value of the potential function does not increase throughout the execution of the algorithm. Thus, if the initial value is K, then K remains an upper bound on the value of $f(\boldsymbol{x}, \boldsymbol{y})$ and by Lemma 2.1, $e^{\frac{K}{\rho - n}}$ is an upper bound on $\boldsymbol{x}^T\boldsymbol{y}$. The essential thing to note is that for every j, the value of $x_j y_j$ is bounded throughout by some K'. Actually, we have demonstrated above that the initial value can be chosen as $\rho \ln n$. Therefore, if $\rho = 2n$, we can set $K' = e^{\frac{K}{n}} = n^2$. Thus,

$$\|\boldsymbol{d}\|^2 \geq n \ \inf\{\lambda(\boldsymbol{M}, \boldsymbol{Y}^{-1}, \boldsymbol{X}) \ : \ x_j, y_j > 0, x_j y_j \leq n^2\} = n\lambda_1(\boldsymbol{M}) \ .$$

∎

6. The case of a P-matrix

Recall that a P-matrix is one that has positive principal minors. Remark 5.2 provides some motivation for considering the performance of the algorithm on problems with P-matrices. In this section we prove that the algorithm presented above solves the LCP with a P-matrix. The algorithm runs until the value of $\boldsymbol{x}^T\boldsymbol{y}$ decreases below some prescribed $\epsilon > 0$. For problems with rational coefficients, an $\epsilon > 0$ can be determined such that an exact solution can be computed from an ϵ-complementary one. The following proposition is the key to the validity of the algorithm:

Proposition 6.1. *If $\boldsymbol{M}$ is a P-matrix, then for every $\epsilon > 0$ there exists $\delta > 0$ such that for every pair $\boldsymbol{x}, \boldsymbol{y}$ produced by the algorithm, if $\boldsymbol{x}^T\boldsymbol{y} > \epsilon$, then necessarily $x_j, y_j > \delta$ for all j.*

Proof: The proof follows from the fact that the potential function decreases monotonically throughout the execution of the algorithm. In particular, there exists a constant K^* such that $f(\boldsymbol{x}, \boldsymbol{y}) < K^*$ for every iterate $(\boldsymbol{x}, \boldsymbol{y})$. By Lemma 2.1, for every such iterate, $\boldsymbol{x}^T\boldsymbol{y} < e^{\frac{K^*}{\rho - n}}$. However, if $\boldsymbol{M}$ is a P-matrix, the set of pairs $\boldsymbol{x}, \boldsymbol{y}$, such that $\boldsymbol{y} = \boldsymbol{M}\boldsymbol{x} + \boldsymbol{q}$ and $\boldsymbol{x}^T\boldsymbol{y} < t$, is bounded for every t. The latter can be seen as follows. The quantity

$$\gamma(\boldsymbol{M}) = \min_{x \neq O} \max_i \frac{x_i(\boldsymbol{M}\boldsymbol{x})_i}{\|\boldsymbol{x}\|^2}$$

is positive for any P-matrix. Suppose, on the contrary, that the set defined above is unbounded. In particular, the inequalities

$$x_i(\boldsymbol{M}\boldsymbol{x})_i + x_i q_i < t \quad (i = 1, \cdots, n)$$

leave $\boldsymbol{x}$ unbounded. However, when $\boldsymbol{x}$ tends to infinity, these inequalities, written in the form

$$\frac{x_i(\boldsymbol{M}\boldsymbol{x})_i}{\|\boldsymbol{x}\|^2} + \frac{x_i q_i}{\|\boldsymbol{x}\|^2} < \frac{t}{\|\boldsymbol{x}\|^2}$$

give a point $\boldsymbol{x}^* \neq \boldsymbol{0}$ such that

$$\frac{x_i^*(\boldsymbol{M}\boldsymbol{x}^*)_i}{\|\boldsymbol{x}^*\|^2} \leq 0$$

for all i, which contradicts $\gamma(\boldsymbol{M}) > 0$. Now, there exists $\epsilon_1 > 0$ such that $\boldsymbol{x}^T\boldsymbol{y} > \epsilon$ and $f(\boldsymbol{x}, \boldsymbol{y}) < K^*$ imply $x_j y_j > \epsilon_1$. The boundedness proven above implies the claim. ∎

We can now define $\lambda^*(\boldsymbol{M}, \epsilon)$ to be the minimum of $\lambda(\boldsymbol{M}, \boldsymbol{Y}^{-1}, \boldsymbol{X})$ where $\boldsymbol{X}$ and $\boldsymbol{Y}$ are such that both $\boldsymbol{e}^T\boldsymbol{X}\boldsymbol{Y}\boldsymbol{e} \geq \epsilon$ and $X_{jj}Y_{jj} \leq n^2$. Since $\boldsymbol{X}$ and $\boldsymbol{Y}$ vary in a compact set, the minimum is attained in the set. In other words, $\lambda^*(\boldsymbol{M}, \epsilon)$ is positive if $\boldsymbol{M}$ is a P-matrix, and this implies that the algorithm computes an ϵ-complementary solution in a finite number of steps. To conclude this section, we give an upper bound on this number. Assume $\rho = 2n$. The initial value of the potential function is $O(n \log n)$. To guarantee ϵ-complementarity, this value has to be reduced to $-n \log \epsilon$. During each step, the value is reduced at least by $\Omega(\|\boldsymbol{d}\|^2/\rho)$ (see Lemma 4.9), which is $\Omega(\lambda^*(\boldsymbol{M}, \epsilon))$ (see Proposition 5.3). Thus, the number of iterations is $O(n \log(n/\epsilon)/\lambda^*(\boldsymbol{M}, \epsilon))$.

7. The positive semi-definite case

In this section we show that the algorithm of this paper solves linear complementarity problems with positive semi-definite matrices in polynomial time. We derive the formulas in the transformed space, so the reader who wishes to use the original matrix $\boldsymbol{M}$, has to replace $\boldsymbol{M}$ in the formulas below by $\boldsymbol{Y}^{-1}\boldsymbol{M}\boldsymbol{X}$. As in Section 4, consider the potential function:

$$f(\boldsymbol{x},\boldsymbol{y}) = \rho \ln \boldsymbol{x}^T\boldsymbol{y} - \sum_{j=1}^{n} \ln x_j y_j$$

where $\rho > n$. The vector $\boldsymbol{g}$ can be written as

$$\boldsymbol{g} = \frac{\rho}{\boldsymbol{x}^T\boldsymbol{y}}\boldsymbol{X}\boldsymbol{Y}\boldsymbol{e} - \boldsymbol{e}\ .$$

The optimality conditions in the problem of projecting $\boldsymbol{h} = (\boldsymbol{g}^T, \boldsymbol{g}^T)^T$ into $\{\boldsymbol{\Delta y} = \boldsymbol{M}\boldsymbol{\Delta x}\}$ are:

$$\begin{aligned}\boldsymbol{\Delta x} - \boldsymbol{g} &= \boldsymbol{M}^T\boldsymbol{\lambda}\\ \boldsymbol{\Delta y} - \boldsymbol{g} &= -\boldsymbol{\lambda}\\ \boldsymbol{\Delta y} &= \boldsymbol{M}\boldsymbol{\Delta x}\ .\end{aligned}$$

Thus we have

$$\begin{aligned}\boldsymbol{\lambda} &= (\boldsymbol{I} + \boldsymbol{M}\boldsymbol{M}^T)^{-1}(\boldsymbol{I} - \boldsymbol{M})\boldsymbol{g}\\ \boldsymbol{\Delta x} &= \boldsymbol{g} + \boldsymbol{M}^T\boldsymbol{\lambda}\\ \boldsymbol{\Delta y} &= \boldsymbol{g} - \boldsymbol{\lambda}\ .\end{aligned}$$

Since the derivation is done in the transformed space, when we use the original matrix $\boldsymbol{M}$ we have to write as follows.

$$\begin{aligned}\boldsymbol{\Delta x} =& \boldsymbol{g} + \frac{\rho}{\boldsymbol{x}^T\boldsymbol{y}}\boldsymbol{X}\boldsymbol{M}^T\boldsymbol{\lambda}'\\ \boldsymbol{\Delta y} =& \boldsymbol{g} - \frac{\rho}{\boldsymbol{x}^T\boldsymbol{y}}\boldsymbol{Y}\boldsymbol{\lambda}'\end{aligned}$$

where

$$\boldsymbol{\lambda}' = \frac{\boldsymbol{x}^T\boldsymbol{y}}{\rho}\boldsymbol{Y}^{-1}(\boldsymbol{I} + \boldsymbol{Y}^{-1}\boldsymbol{M}\boldsymbol{X}^2\boldsymbol{M}^T\boldsymbol{Y}^{-1})^{-1}(\boldsymbol{I} - \boldsymbol{Y}^{-1}\boldsymbol{M}\boldsymbol{X})\boldsymbol{g}.$$

Thus,

$$\begin{aligned}\boldsymbol{\Delta x} =& \frac{\rho}{\boldsymbol{x}^T\boldsymbol{y}}\boldsymbol{X}(\boldsymbol{y} + \boldsymbol{M}^T\boldsymbol{\lambda}') - \boldsymbol{e}\\ \boldsymbol{\Delta y} =& \frac{\rho}{\boldsymbol{x}^T\boldsymbol{y}}\boldsymbol{Y}(\boldsymbol{x} - \boldsymbol{\lambda}') - \boldsymbol{e}\ .\end{aligned}$$

We can now prove the following:

Proposition 7.1. *If $\boldsymbol{M}$ is positive semi-definite and $\rho = 2n + \sqrt{2n}$, then $\|\boldsymbol{d}\| \geq 1$.*

Proof: Denote

$$\begin{aligned}\boldsymbol{x}' =& \boldsymbol{x} - \boldsymbol{\lambda}'\\ \boldsymbol{y}' =& \boldsymbol{y} + \boldsymbol{M}^T\boldsymbol{\lambda}'\ .\end{aligned}$$

We have

$$\begin{aligned}\boldsymbol{\Delta x} &= \frac{\rho}{\boldsymbol{x}^T\boldsymbol{y}}\boldsymbol{X}\boldsymbol{y}' - \boldsymbol{e}\\ \boldsymbol{\Delta y} &= \frac{\rho}{\boldsymbol{x}^T\boldsymbol{y}}\boldsymbol{Y}\boldsymbol{x}' - \boldsymbol{e}\ .\end{aligned}$$

If either $\boldsymbol{x}' \not\geq \boldsymbol{0}$ or $\boldsymbol{y}' \not\geq \boldsymbol{0}$, then at least one coordinate of $\boldsymbol{d} = (\boldsymbol{\Delta x}, \boldsymbol{\Delta y})$ is less than -1, so $\|\boldsymbol{d}\| \geq 1$. Thus, assume $\boldsymbol{x}', \boldsymbol{y}' \geq \boldsymbol{0}$. Also,

$$(\boldsymbol{y}' - \boldsymbol{y})^T(\boldsymbol{x}' - \boldsymbol{x}) = -(\boldsymbol{\lambda}')^T\boldsymbol{M}\boldsymbol{\lambda}' \leq 0\ .$$

Thus,

$$\boldsymbol{x}^T\boldsymbol{y}' + \boldsymbol{y}^T\boldsymbol{x}' \geq \boldsymbol{x}^T\boldsymbol{y} + \boldsymbol{x}'^T\boldsymbol{y}' \geq \boldsymbol{x}^T\boldsymbol{y}\ > 0.$$

This implies

$$\begin{aligned}\|\boldsymbol{d}\|^2 &= \frac{\rho^2}{(\boldsymbol{x}^T\boldsymbol{y})^2}\left(\|\boldsymbol{X}\boldsymbol{y}'\|^2 + \|\boldsymbol{Y}\boldsymbol{x}'\|^2\right) - 2\frac{\rho}{\boldsymbol{x}^T\boldsymbol{y}}(\boldsymbol{x}^T\boldsymbol{y}' + \boldsymbol{y}^T\boldsymbol{x}') + 2n\\ &\geq \frac{\rho^2}{(\boldsymbol{x}^T\boldsymbol{y})^2}\frac{(\boldsymbol{x}^T\boldsymbol{y}')^2 + (\boldsymbol{y}^T\boldsymbol{x}')^2}{n} - 2\frac{\rho}{\boldsymbol{x}^T\boldsymbol{y}}(\boldsymbol{x}^T\boldsymbol{y}' + \boldsymbol{y}^T\boldsymbol{x}') + 2n\\ &\geq \frac{\rho^2}{(\boldsymbol{x}^T\boldsymbol{y})^2}\frac{(\boldsymbol{x}^T\boldsymbol{y}' + \boldsymbol{y}^T\boldsymbol{x}')^2}{2n} - 2\frac{\rho}{\boldsymbol{x}^T\boldsymbol{y}}(\boldsymbol{x}^T\boldsymbol{y}' + \boldsymbol{y}^T\boldsymbol{x}') + 2n\\ &= 2n\left(\frac{\rho(\boldsymbol{x}^T\boldsymbol{y}' + \boldsymbol{y}^T\boldsymbol{x}')}{2n\boldsymbol{x}^T\boldsymbol{y}} - 1\right)^2\\ &\geq 2n\left(\frac{\rho}{2n} - 1\right)^2\\ &= 2n\left(1 + \frac{1}{\sqrt{2n}} - 1\right)^2 = 1\ .\end{aligned}$$

∎

Finally, we can state the following:

Theorem 7.2. *If $\boldsymbol{M}$ is positive semi-definite then, starting at any interior point where the potential value is $O(nL)$, the algorithm converges in $O(n^2L)$ iterations.*

Proof: If $\boldsymbol{M}$ is positive semi-definite, it follows from Proposition 7.1 and Lemma 4.9 (with $\rho = 2n + \sqrt{2n}$) that

$$\Delta f \geq \frac{1}{O(n)}\ .$$

Thus, by Lemma 2.1, after $O(n^2L)$ iterations we have $\boldsymbol{x}^T\boldsymbol{y} \leq 2^{-L}$. ∎

Initializations that fit the above theorem and the next one were given elsewhere (see, for example, [4]).

Theorem 7.3. *If $\boldsymbol{M}$ is skew-symmetric (including the case of the linear programming problem) then, starting at any interior point where the potential value is $O(nL)$, the algorithm converges in $O(nL)$ iterations.*

Proof: In the skew-symmetric case we obtain better estimates as follows. First, in Proposition 4.5, the quadratic term vanishes, so

$$\ln(\boldsymbol{e}^T\boldsymbol{W}\boldsymbol{e}) - \ln\left((\boldsymbol{e} - t\boldsymbol{\Delta x})^T\boldsymbol{W}(\boldsymbol{e} - t\boldsymbol{\Delta y})\right) \geq \frac{\boldsymbol{e}^T\boldsymbol{W}(\boldsymbol{\Delta x} + \boldsymbol{\Delta y})}{\boldsymbol{e}^T\boldsymbol{W}\boldsymbol{e}}\, t \ .$$

Hence, the inequality of Corollary 4.8 becomes

$$\Delta f \geq \|\boldsymbol{d}\|^2 t - 2\|\boldsymbol{d}\|^2 t^2 \ .$$

Following arguments similar to the ones used in the proof of Lemma 4.9, let $t' = 1/4$ denote the maximizer of the parabola in the right-hand side. If $t' < t^*$, then we have $\Delta f \geq \|\boldsymbol{d}\|^2/8$. Otherwise, we get $\Delta f \geq \|\boldsymbol{d}\|/4$, and the claim follows from Proposition 7.1. ∎

References

[1] D. Gale and H. Nikaido, "The Jacobian matrix and global univalence of mappings", *Mathematische Annalen* **159** (1965) 81–93.

[2] C. C. Gonzaga, "Polynomial affine algorithms for linear programming," unpublished manuscript, Federal University of Rio de Janeiro, 1988.

[3] N. Karmarkar, "A new polynomial-time algorithm for linear programming", *Combinatorica* **4** (1984) 373–395.

[4] M. Kojima, S. Mizuno and A. Yoshise, "Polynomial time algorithm for a class of linear complementarity problems", *Mathematical Programming* **44** (1989) 1–26.

[5] N. Megiddo, "A Note on the Complexity of P-Matrix LCP and Computing an Equilibrium", Research Report RJ 6439, IBM Almaden Research Center, San Jose, CA 95120-6099, 1988.

[6] Y. Ye, "An $O(n^3L)$ potential reduction algorithm for linear programming," unpublished manuscript, The University of Iowa, 1989.