Fall 2012

Infeasible Full-Newton-Step Interior-Point Method for the Linear Complementarity Problems

Antré Marquel Drummer

INFEASIBLE FULL-NEWTON-STEP INTERIOR-POINT METHOD FOR THE LINEAR COMPLEMENTARITY PROBLEMS

by

ANTRE' MARQUEL DRUMMER

(Under the Direction of Goran Lesaja)

ABSTRACT

In this thesis, we present a new Infeasible Interior-Point Method (IPM) for monotone Linear Complementarity Problem (LPC). The advantage of the method is that it uses full Newton-steps, thus, avoiding the calculation of the step size at each iteration. However, by suitable choice of parameters the iterates are forced to stay in the neighborhood of the central path, hence, still guaranteeing the global convergence of the method under strict feasibility assumption. The number of iterations necessary to find ϵ-approximate solution of the problem matches the best known iteration bounds for these types of methods. The preliminary implementation of the method and numerical results indicate robustness and practical validity of the method.

INDEX WORDS: linear complementarity problem, interior-point method, full Newton-step, polynomial convergence

INFEASIBLE FULL-NEWTON-STEP INTERIOR-POINT METHOD FOR THE LINEAR COMPLEMENTARITY PROBLEMS

by

ANTRE' MARQUEL DRUMMER

B.S special degree in Mathematics, Georgia Southern University, 2010

A Thesis Submitted to the Graduate Faculty of Georgia Southern University in Partial Fulfillment of the Requirement for the Degree

MASTER OF SCIENCE

STATESBORO, GEORGIA

2012

INFEASIBLE FULL-NEWTON-STEP INTERIOR-POINT METHOD FOR THE LINEAR COMPLEMENTARITY PROBLEMS

by

ANTRE' MARQUEL DRUMMER

Major Professor:	Goran Lesaja
Committee:	Scott Kersey
	Frederick Mynard

Electronic Version Approved:

June, 2012

DEDICATION

This thesis is dedicated to all who believed in and supported me on this journey. I'm truly greatful for all the prayers and encouragement.

ACKNOWLEDGMENTS

I thank God for all of his many blessings, and among those is Dr. Goran Lesaja. Words can not begin to express how greatful I am to Dr. Goran Lesaja. His patience, guidance, and knowledge led me through this thesis. The success of this work is credited to Dr. Lesaja's exceptional direction. It was a privilege and an honor to work under a man such as he. Not only is Dr. Lesaja a great advisor, but I also found him to be a great friend. Also, special thanks goes to my committee members. I also acknowledge every member of the outstanding faculty and staff in Department of Mathematical Siences at Georgia Southern University, especially Dr. Martha Abell, Department Chair, and Dr. Yan Wu, Graduate Program Director, for all their utmost commitments toward the success of the students and the department.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Problem

In this thesis we shall consider a Linear Complementarity Problem (LCP) in the standard form:

$$f(x) = s, \quad x \geq 0, \quad s \geq 0, \quad x^T s = 0 \tag{1.1}$$

where $x, s \in R^n$ and f is a linear function

$$f(x) = Mx + q \geq 0$$

where matrix $M \in R^{n \times n}$ and $q \in R^n$ are given. In other words, the objective of LCP is to find nonegative vectors (x, s) that satisfy the linear equation $s = Mx + q$ and are orthogonal i.e. $x^T s = 0$. Though the LCP is not an optimization problem, there are a plethora of optimization problems that can be modeled as LCP directly or indirectly. Some applications of LCP in operations research include but are not limited to game theory, economics, and many more.

The relationship between LCP and optimization problems is very close because Kurush-Kuhn-Tucker (KKT) optimality conditions for many optimization problems can be formulated as LCP. We consider the connection between LCP, an example of which will be given in Example 1 of Chapter 2, and the linear programming problem (LP). The LP can be formulated as a maximization or minimization problem which has the form

$$\begin{aligned} \min \quad & c^T x \geq 0 \\ \text{s.t.} \; & Ax = b \\ & x \geq 0 \end{aligned} \tag{1.2}$$

where $c \in R^n,\ b \in R^m,\ and\ A \in R^{n\times n}$. Similar to the objective for LCP, the objective of the LP is to find the vector x that satisfy the equations and constratints of (1.2).

1.2 A Brief Historical Overview

Some instances of the LCP can be traced back to the early 1940's; however, larger interest in LCP was taken in the early to mid 1960's in conjunction with the rapid development of theory and methods for LP.

In 1947, George Dantzig proposed a famous method, named the Simplex Method (SM) for solving the LP. Basically, the main idea of the SM is to travel along from vertex to vertex on the boundary of the feasible region. The method constantly increases (or decreases) the objective function until either an optimal solution is found or the SM concludes that such an optimal solution does not exist.

Theoretically, the algorithm could have a worse-case scenario of 2^n iteration, with n being the size of the problem, which is an exponential number. This was shown in 1972 by Klee and Minty [8]. However, it is remarkably efficient in practice but an exponential number of iterations is usually never observed in practice. It usually requires $O(n)$ iterations to solve a particular problem. There exists many resources and excellent software for the SM.

Another great advancement in the area of solving convex optimization problems was the *ellipsoid method*. This method was introduced by Nemirovsky and Yudin [24] in 1976 and by Shor [20] in 1977. The algorithm works by encapsulating the minimizer of a convex function in a sequence of ellipsoids whose volume decreases at each iteration. Later Khachiyan [7] showed in 1984 that the ellipsoid method can be used to solve the LP in polynomial time. This was the first polynomial time algorithm for the LP. Unfortunately, in practice, the method was far surpassed by the SM. Nevertheless, the theoretical importance of the ellipsoid method is hard to neglect.

In 1984, Karmarkar [6] introduced an *Interior-Point Method* (IPM) for LP. Karmarkar used the efficiency of the simplex method with the theoretical advantages of the ellipsoid method to create his efficient polynomial algorithm. The algorithm is based on projective transformations and the use of Karmarkar's primal potential function. This new algorithm sparked much research, creating a new direction in optimization - the field of IPMs. Unlike the SM, which travels from vertex to vertex along the edges of the feasible region, the IPM follows approximately a central path in the interior of the feasible region and reaches the optimal solution only asymptotically. As a result of finding the optimal solution in this fashion, the analysis of the IPMs become substantially more complex than that of the SM.

Since the first IPM was developed, many new and efficient IPMs for solving LP have been created. Many researchers have proposed different interior-point methods, which can be grouped into two different groups: potential reduction algorithms and path-following algorithms. Each of the two groups contains algorithms based on primal, dual, or primal-dual formulations of the LP. Also, computational results show that the primal-dual formulation is superior to either the primal or dual formulation of the algorithm. We will focus on the *primal-dual* path-following IPMs, which have become the standard of efficiency in practical applications. These primal-dual methods are based on using Newton's method in a careful and controlled manner.

Soon after the SM was developed, a similar method for solving LCP was introduced by Lemke [10]. It is a pivoting algorithm similar to the SM. Unfortunately, Lemke's algorithm can sometimes fail to produce a solution even if one exists. Nevertheless, Lemke's algorithm was extremely useful. However, researchers kept searching for other methods for the LCP. Much later, in the 1990's, the tradition of immediate generalizations from LP to LCP continued even more strongly in the case of the IPMs and many efficient IPMs have been proposed for LCP.

In this thesis, we will focus on extending a class of IPMs, from LP to LCP. The main features of this class of methods is that at each iteration a full Newton-step is taken, i.e., it is not necessary to calculate a step size. These type of IPMs are called Full-Newton-step IPM (FNS-IPM). They were first discussed for LP by Roos in [18].

In addition, IPMs have been generalized to solve many other important optimization problems, such as semidefinite optimization, second order cone optimization, and general convex optimization problems. The unified theory of IPMs for general convex optimization problems was first developed by Nesterov and Nemirovski [15] in 1994.

The first comprehensive monograph that considers in-depth analysis of the LCP and methods for solving it is the monograph of Cottle, Pang, and Stone [3]. More recent results on the LCP as well as nonlinear complementarity problems and variational inequalities are contained in the monograph of Facchinei and Pang [5].

CHAPTER 2

LINEAR COMPLEMENTARITY PROBLEM

Chapter 2 shall consist of the introduction of the linear complementarity problem (LCP). Along with the discussion and defining of the LCP, several direct applications will also be presented and discussed.

2.1 Linear Complementarity Problem

LCP is a problem of finding a particular vector in a finite real vector space that satisfies a system of inequalities. In mathematical terminology this means: given a vector $q \in \mathbb{R}^n$ and a matrix $M \in \mathbb{R}^{nxn}$, we want to find a pair of vectors $x, s \in \mathbb{R}^n$ (or show such a vector does not exist) such that

$$\begin{aligned} s &= q + Mx \\ x \geq 0, s &\geq 0 \\ x^T s &= 0. \end{aligned} \tag{2.1}$$

To insure a solution exists and it is unique, a sufficient condition is that M be a symmetric positive definite matrix. Since $(x, s) \geq 0$, the complementarity equation $x^T s = 0$ can be written equivalently as

$$xs = 0,$$

which represents component-wise product of vectors, as follows,

$$xs = (x_1 s_1, x_2 s_2, \ldots, x_n s_n)^T. \tag{2.2}$$

This product is termed as Hadamard's product.

The feasible region (set of feasible points) of the LCP as defined in (2.1) is given in the following set:

$$F = \left\{(x, s) \in \mathbb{R}^{2n} : s = Mx + q, x \geq 0, s \geq 0\right\}. \tag{2.3}$$

Furthermore, the set of strictly feasible points of the LCP is the following set:

$$F_0 = \{(x, s) \in F : x > 0, s > 0\}.$$

The solution set of the LCP is given by

$$F^* = \left\{(x^*, s^*) \in F : x^{*T} s^* = 0\right\}. \tag{2.4}$$

An important subset of the above solution set is a set of strict complementarity solutions

$$F_s^* = \{(x^*, s^*) \in F^* : x^* + s^* > 0\}. \tag{2.5}$$

We can now say that the main idea of the LCP is to find vectors x, s (a solution of the LCP) that are both feasible and complementary. If $q \geq 0$, the LCP is always solvable with the zero vector being a trivial solution.

2.2 Classes of LCP

In general LCP is NP-complete, which means that there exists no polynomial algorithms for solving it. Thus, the problem needs to be restricted to certain classes of matrices for which a polynomial algorithm exist. We now list several such classes of matrices M for LCP. They are as follows:

- Skew-symmetric matrices (SS):

$$(x \in \mathbb{R}^n)(x^T M x = 0). \tag{2.6}$$

- Positive semi-definite matrices (PSD):

$$(x \in \mathbb{R}^n)(x^T Mx \geq 0). \tag{2.7}$$

- P-matrices: Matrices with all principal minors positive or equivalently

$$(0 \neq x \in \mathbb{R}^n)(\exists i \in I)(x_i(Mx)_i > 0). \tag{2.8}$$

- P_0-matrices: Matrices with all principal minors nonnegative or equivalently

$$(0 \neq x \in \mathbb{R}^n)(\exists i \in I)(x_i \neq 0 \text{ and } x_i(Mx)_i \geq 0). \tag{2.9}$$

- Sufficient matrices (SU): Matrices which are column and row sufficient
  - Column sufficient matrices (CSU):

$$(\forall x \in \mathbb{R}^n)(\forall i \in I)(x_i(Mx)_i \leq 0 \Rightarrow x_i(Mx)_i = 0). \tag{2.10}$$

  - Row sufficient matrices (RSU): M is row sufficient if M^T is column sufficient.

- $P_*(\kappa)$: Matrices such that

$$(1+4\kappa) \sum_{i \in I^+(x)} x_i(Mx)_i + \sum_{i \in I^-(x)} x_i(Mx)_i \geq 0, \forall x \in \mathbb{R}^n,$$

where

$$I^+(x) = \{i : x_i(Mx)_i > 0\}, I^-(x) = \{i : x_i(Mx)_i < 0\},$$

or equivalently

$$x^T Mx \geq -4\kappa \sum_{i \in I^+(x)} x_i(Mx)_i, \forall x \in \mathbb{R}^n, \tag{2.11}$$

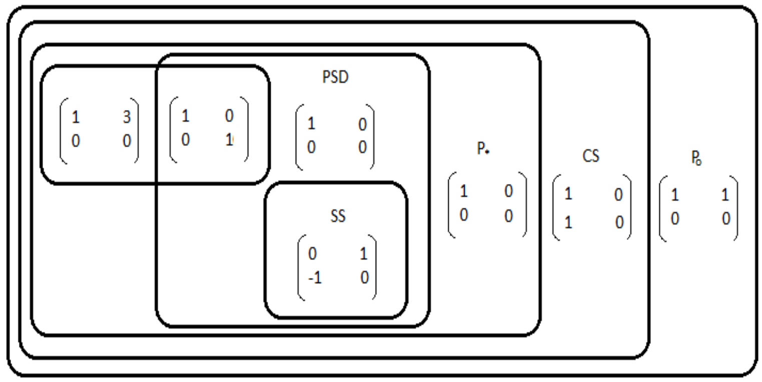

Figure 2.1: Relations and examples of the classes of matrices.

and

$$P_* = \bigcup_{\kappa \geq 0} P_*(\kappa). \tag{2.12}$$

The relationship between some of the above classes is as follows:

$$SS \subset PSD \subset P_* = SU \subset CS \subset P_0, \quad P \subset P_*, \quad P \cap SS = \emptyset. \tag{2.13}$$

Some of these relations are obvious, like PSD $= P_*(0) \subset P_*$ or $P \subset P_*$, while others require proof. Refer to Figure 2.1, which was first published in [9], to see a visual flow of how these classes of matrices are related. Also, all of the above classes have the nice property that if a matrix M belongs to one of these classes, then every principal sub-matrix of M also belongs to the class.

In this thesis, we will assume matrix M is a positive semi-definite (PSD) matrix. Though the case of positive semi definiteness is not the most general case, it is definitely the most commonly used both in theory and practice. Hence, this is the reason why we will focus on this class of matrices in the thesis. The LCP with a PSD matrix M is called *monotone* LCP.

2.3 Introductory Examples

LCP has many applications. Some examples of the LCP include but are by far not limited to: the bimatrix game, optimal invariant capital stock, optimal stopping, convex hulls in the plane, and the market equilibrium problems. Each one of the listed problems can be reformulated into the linear complementarity problem. In the sequel, we will describe several applications.

Example 1: Quadratic Programming

Quadratic programming is another application of the LCP. It is the problem of minimizing or maximizing a quadratic function of several variables subject to linear constraints on these variables. The quadratic program (QP) is defined as

$$\begin{aligned} \text{minimize} \quad & f(x) = c^T x + \tfrac{1}{2} x^T Q x \\ \text{subject to} \quad & Ax \geq b \\ & x \geq 0 \end{aligned} \tag{2.14}$$

where $Q \in \mathbb{R}^{nxn}$ is symmetric, $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{mxn}$ and $b \in \mathbb{R}^m$. Note that the case where $Q = 0$ gives rise to a linear program (LP). If x is a locally optimal solution of the quadratic program (2.14), then there exists a vector $y \in \mathbb{R}^m$ such that the pair (x, y) satisfies the Karush-Kuhn-Tucker optimality conditions

$$\begin{aligned} u &= c + Qx - A^T y \geq 0, \quad x \geq 0, \quad x^T u = 0, \\ v &= -b + Ax \geq 0, \quad\quad\quad\; y \geq 0, \quad y^T v = 0. \end{aligned} \tag{2.15}$$

If Q is positive semi-definite (the objective function $f(x)$ is convex), then the conditions in (2.15) are sufficient for the vector x to be a globally optimal solution of (2.14).

The Karush-Kuhn-Tucker conditions in (2.14) define the LCP where

$$q = \begin{bmatrix} c \\ -b \end{bmatrix} \quad \text{and} \quad M = \begin{bmatrix} Q & -A^T \\ A & 0 \end{bmatrix}. \tag{2.16}$$

Note that M is not symmetric, even though Q is symmetric. However, M does have a property known as *bisymmetry*. A square matrix A is *bisymmetric* if it can be brought to the form

$$A = \begin{bmatrix} G & -A^T \\ A & H \end{bmatrix},$$

where both G and H are symmetric. Also, if Q is positive semi-definite, then so is M. In general, a square matrix M is positive semi-definite if $z^T M z \geq 0$ for every vector z.

This convex quadratic programming model, in the form of (2.14), has a magnitude of practical applications in engineering, finance, and many other areas. The size of these practical problems can become very large. Thus, the LCP plays an important role in the numerical solution of these problems.

Example 2: Bimatrix games

Game theory analyzes strategic interactions in which the outcome of one's choices depends upon the choices of others. For a situation to be considered a game, there must be at least two rational players who take into account one another's actions when formulating their own strategies. We consider a game with two players called player I and player II and the game consists of large number of plays. Here at each play Player I picks one of m choices and Player II picks one of n choices. These choices are called pure strategies. If in a certain play, Player I choose pure strategy i and Player II chooses pure strategy j, then Player I loses A_{ij} and Player II loses B_{ij}. A positive value of A_{ij} represents a loss to Player I, while a negative value of A_{ij} represents a gain. Similarly for Player II and B_{ij}. The matrices A and B are called loss matrices, and the game is fully determined by the matrix pair (A, B).

If $A + B = 0$, the game is known as zero sum game and if $A + B \neq 0$ game is known as bimatrix game. Player I chooses to play strategy i with probability x_i such that $\sum x_i = 1$, and Player II chooses to play strategy j with probability y_j such that $\sum y_j = 1$, then expected loss of Player I is $x^T Ay$ and expected loss of Player II is $x^T By$.

A player is changing his own strategy while the other player holds his strategy fixed to minimize loss. i.e,

$$\begin{aligned} \overline{x}^T A\overline{y} &\leq x^T A\overline{y} \quad \forall x \geq 0 \quad e_m^T x = 1, \\ \overline{x}^T B\overline{y} &\leq \overline{x}^T By \quad \forall y \geq 0 \quad e_n^T y = 1, \end{aligned} \tag{2.17}$$

where the vector e is a vector of ones. The objective is to find $(\overline{x}, \overline{y})$ that is called Nash equilibrium pair. Nash equilibrium can be found using LCP as described in the Lemma below.

Lemma 2.3.1. *Suppose $A, B \in \mathbb{R}^{m\times n}$ are positive loss matrices representing a game $\Gamma(A, B)$ and suppose that $(s, t) \in \mathbb{R}^{m\times n}$ solves LCP(M,q), where*

$$M = \begin{bmatrix} 0 & A \\ B^T & 0 \end{bmatrix}, q = -e_{m+n} \in \mathbb{R}^{m+n}.$$

Then $(\overline{x}, \overline{y})$ *such that,*

$$\overline{x} = \tfrac{s}{e_m^T s} \text{ and } \overline{y} = \tfrac{t}{e_m^T t},$$

is an equilibrium pair of $\Gamma(A, B)$*.*

Proof. We write LCP conditions explicitly as

$$\begin{aligned} 0 \leq At - e_m \perp s \geq 0 \\ 0 \leq B^T s - e_n \perp t \geq 0 \end{aligned} \tag{2.18}$$

from the equation (2.1) we have $Mx + q = s \geq 0$ and $x \geq 0$. So we can write these inequalities as below,

$$\begin{bmatrix} 0 & A \\ B^T & 0 \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix} + \begin{bmatrix} e_m \\ e_n \end{bmatrix} \geq 0, \quad \begin{bmatrix} At \\ B^T s \end{bmatrix} + \begin{bmatrix} e_m \\ e_n \end{bmatrix} \geq 0. \tag{2.19}$$

This implies $At - e_m \geq 0$ and $B^T s - e_n \geq 0$. Therefore $t \neq 0$ and $s \neq 0$. Then $\overline{x} = \frac{s}{e_m^T s}$ and $\overline{y} = \frac{t}{e_n^T t}$ well define. $\overline{x} \geq 0$, $\overline{y} \geq 0$, from the definition we have $e_m^T \overline{x} = 1$ and $e_n^T \overline{y} = 1$. Then $\overline{x}$ and $\overline{y}$ are mixed strategies. By complementarity we have,

$$\overline{x}^T (At - e_m) = \frac{s^T}{e_m^T s}(At - e_m) = 0. \tag{2.20}$$

Since $\overline{x}$ and $\overline{y}$ are mixed strategies, and from the Equation (2.20), we get the following property.

$$\overline{x}^T At = \overline{x}^T e_m = 1. \tag{2.21}$$

So we have,

$$\begin{aligned} A\overline{y} - (\overline{x}^T A\overline{y})e_m &= \tfrac{1}{e_n^T t}(At) - (\overline{x}^T A\overline{y})e_m \\ &= \tfrac{1}{e_n^T t}(At - (\overline{x}^T At)e_m) \\ &= \tfrac{1}{e_n^T t}(At - e_m) \quad \text{from} \quad (2.21) \end{aligned} \tag{2.22}$$

Since $At - e_m \geq 0$ and $x \geq 0$, we have $x^T(A\overline{y} - (\overline{x}^T A\overline{y})e_m) \geq 0$. This implies,

$$x^T A\overline{y} \geq (x^T e_m)(\overline{x}^T A\overline{y}) = \overline{x}^T A\overline{y} \tag{2.23}$$

Similarly we can prove $\overline{x}^T By \geq \overline{x}^T B\overline{y}$. Hence $(\overline{x}, \overline{y})$ is a Nash equilibrium pair. □

Example 3: The Market Equilibrium Problem

The state of an economy where the supplies of producers and the demands of consumers are balanced at the resulting price level is called *market equilibrium* . We can use a linear programming model to describe the supply side that captures technological details of production activities for a particular market equilibrium problem. Econometric models with commodity prices as the primary independent variables generates the market demand function. Basically, we need to find a vector x^* and subsequent vectors p^* and r^* such that the conditions below are satisfied for supply, demand, and equilibrium:

Supply conditions:

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax \geq b \\ & Bx \geq r^* \\ & x \geq 0 \end{aligned} \tag{2.24}$$

where c is the cost vector for the supply activities, x is the vector production activities. Technological constraints on production are represented by the first condition in (2.24) and the demand requirement constraints are represented by the second condition in (2.24);

Demand conditions:

$$r^* = Q(p^*) = Dp^* + d \tag{2.25}$$

where $Q(\cdot)$ is the market demand function with p^* and r^* representing the vectors of demand prices and quantities, respectively. $Q(\cdot)$ is assumed to be an affine function;

Equilibrium condition:

$$p^* = \pi^* \tag{2.26}$$

where the (dual) vector of market supply prices corresponding to the second constraint in (2.24) is denoted by π^*.

Using Karush-Kuhn-Tucker conditions for problem (2.24), we see that a vector x^* is an optimal solution of problem (2.24) if and only if there exists vectors v^* and π^* such that:

$$\begin{aligned} y^* &= c - A^T v^* - B^T \pi^* \geq 0, & x^* \geq 0, \quad (y^*)^T x^* = 0, \\ u^* &= -b + Ax^* \geq 0, & v^* \geq 0, \quad (u^*)^T v^* = 0, \\ \delta^* &= -r^* + Bx^* \geq 0, & \pi^* \geq 0, \quad (\delta^*)^T \pi^* = 0. \end{aligned} \tag{2.27}$$

If for r^*, we substitute the demand function (2.25) and we use condition (2.26), then we can see that the conditions in (2.27) gives us the linear complementarity problem where

$$q = \begin{bmatrix} c \\ -b \\ -d \end{bmatrix}, \ M = \begin{bmatrix} 0 & -A^T & -B^T \\ A & 0 & 0 \\ B & 0 & -D \end{bmatrix}. \tag{2.28}$$

Observe that the matrix M in (2.28) is bisymmetric and if the matrix D is symmetric, as it could have been seen, the Karush-Kuhn-Tucker optimization conditions of the market equilibrium problem, and in fact the linear problem in general, can be expressed in the LCP framework. This can also be extended to quadratic programming problems as discussed below.

$$\begin{aligned} \text{maximize} \quad & d^T x + \tfrac{1}{2} x^T D x + b^T y \\ \text{subject to} \quad & A^T y + B^T x \leq c \\ & x \geq 0, \quad y \geq 0 \end{aligned} \tag{2.29}$$

On the other hand, if D is asymmetric, then M is not bisymmetric and the connection between the market equilibrium model and the quadratic program above fails to exist.

CHAPTER 3

LEMKE'S METHOD

In this chapter, we review a well known algorithm called Lemke's algorithm. Lemke's Method, derived in 1965, is the first algorithm proposed for solving LCPs. This is a pivoting algorithm introduced by Lemke [10] and it is a generalization of Dantzig's Simplex Method developed earlier for LP.

3.1 Basic Definition

We consider a LCP in the standard form as described in (2.1) Chapter 2. We denote it here as LCP(M,q) and claim that (x, s) is *feasible* for LCP(M,q) if all conditions of the following system are satisfied.

$$\begin{aligned} s &= q + Mx \\ x &\geq 0, s \geq 0. \end{aligned} \tag{3.1}$$

Proceeding, we assume that M is a positive semidefinite (psd) matrix. For the description of Lemke's method for solving LCP(M,q), we introduce the following definitions.

Definition 3.1.1.

Consider the problem LCP(M,q) (3.1).

1. A component s_i is called the complement of x_i, and vice versa, for $i = 1, 2, ..., n$.
2. A pair (x, s) is complementary if $x \geq 0, s \geq 0$, and $x^T s = 0$. (Note that a complementary pair must satisfy $x_i s_i = 0$ for $i = 1, 2, ..., n$.)

3. A pair (x, s) is almost complementary if $x \geq 0, s \geq 0$, and $x_i s_i = 0$ for $i = 1, 2, ..., n$, except for a single index j, $1 \leq j \leq n$.

3.2 Lemke's Method

For a positive semidefinite matrix M, Lemke's method generates a finite sequence of feasible, almost-complementary pairs that terminates at a complementary pair or an unbounded ray. That is, for any feasible solution x with objective, a multiple of the unbounded ray can be added to x to give a feasible solution with objective z-1 (or z+1 for maximization models). Thus, if a feasible solution exists, then the optimal objective is unbounded.

Similar to the Simplex Method, an initial pair must first be obtained, usually via a Phase I scheme. There are different Phase I schemes depending on the particular structure of LCP. We will describe a commonly used Phase I scheme, which requires only one pivot.

Phase II generates a sequence of almost-complementary vector pairs. It performs a pivot at each iteration, selecting the pivot row by means of a ratio test like that of the Simplex Method, whose purpose is to ensure that the components of x and s remain nonnegative throughout the procedure. Phase II finishes when the complementary pair is found or we end up on the unbounded ray.

This outline can be summarized as follows.

Lemke's Algorithm

Phase I: (Generates a Feasible Almost- Complementary Table).

1. If $q \geq 0$, STOP : $x = 0$ is a solution of LCP(M,q); that is, $(x, s) = (0, q)$ is a feasible complementary pair.

2. Otherwise, add the artificial variables x_0 and s_0 that satisfy the following relationships:

$$s = Mx + ex_0 + q, s_0 = x_0, \tag{3.2}$$

where e is the vector of ones in $\mathbb{R}^n$. Create the initial tableau,

	x	x_0	1
$s =$	M	e	q
$s_0 =$	0	1	0

3. Make this tableau feasible by carrying out a Jordan exchange on the x_0 column and the row corresponding to the most negative q_i.

4. Without removing the artificial variables from the tableau, proceed to *Phase II*.

(**Phase II**: Generates a Feasible Complementary or Unbounded Tableau).

1. Start with a feasible almost-complementary pair (x, s) and the corresponding tableau in Jordan exchange form,

	s_{I_1}	x_{J_2}	1
$x_{J_1} =$	$H_{I_1 J_1}$	$H_{I_1 J_2}$	h_{I_1}
$s_{I_2} =$	$H_{I_2 J_1}$	$H_{I_2 J_2}$	h_{I_2}

Record the variable that became nonbasic (i.e., became a column label) at the previous iteration. At the first step, this is simply the component of s that was exchanged with x_0 during Phase I.

2. Pivot column selection: Choose the column s corresponding to the complement of the variable that became nonbasic at the previous pivot.

3. Pivot row selection: Choose the row r such that,

$$-h_r/H_{rs} = min\,\{-h_i/H_{is}|H_{is} < 0\}.$$

   If all $H_{is} \geq 0$, STOP: An unbounded ray has been found.

4. Carry out a Jordan exchange on element H_{rs}. If (x, s) is complementary, STOP: (x, s) is a solution. Otherwise, go to Step 2.

Remarks

1. Step 2 maintains almost-complementarity by moving a component into the basis as soon as its complement is moved out. By doing so, we ensure that for all except one of the components, exactly one of x_i and s_i is basic while the other is nonbasic. Since nonbasic variables are assigned the value 0, this fact ensures that $x_i s_i = 0$ for all except one component which is the almost complementary property. When the initial tableau of Phase II was derived from Phase I, it is the variables s_0 and x_0 that violate complementarity until an optimal tableau is found.

2. The ratio test in Step 3 follows from the same logic as in the Simplex Method. We wish to maintain non negativity of all the components in the last column, and so we allow the nonbasic variable in column s to increase away from zero only until it causes one of the basic variables to become zero. This basic variable is identified by the ratio test as the one to leave the basis in the current iteration.

3. In practice, it is not necessary to insert the s_0 row into the tableau, since s_0 remains in the basis throughout and is always equal to x_0.

The following important theorem assures that Lemke's algorithm terminates at the solution of the LCP(M,q) if M is positive semidefinite.

Theorem 3.2.1. *1. If $M \in \mathbb{R}^{n\times n}$ is positive definite, then Lemke's algorithm terminates at the unique solution of LCP(M,q) for any $q \in \mathbb{R}^n$.*

2. If $M \in \mathbb{R}^{n\times n}$ is positive semidefinite, then for each $q \in \mathbb{R}^n$, Lemke's algorithm terminates at a solution of LCP(M,q) or at an unbounded ray. In the latter case, the set $\{x|Mx + s \geq 0, x \geq 0\}$ is empty; that is, there is no feasible pair.

The proof can be found in [4].

3.3 Example

We consider a quadratic programming problem

$$\begin{aligned} \min \quad & \tfrac{1}{2}x_1^2 - x_1x_2 + \tfrac{1}{2}x_2^2 + 4x_1 - x_2 \\ \text{s.t.} \quad & x_1 + x_2 - 2 \geq 0 \\ & x_1, x_2 \geq 0. \end{aligned} \tag{3.3}$$

The KKT condition of this problem are described in Example 1, Chapter 2, (2.15) and (2.16). In this case we have

$$Q = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \; A = \begin{bmatrix} 1 & 1 \end{bmatrix}, \; p = \begin{bmatrix} 4 \\ -1 \end{bmatrix}, \; b = \begin{bmatrix} 2 \end{bmatrix},$$

which leads to the following LCP

$$M = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ 1 & 1 & 0 \end{bmatrix}, \ q = \begin{bmatrix} 4 \\ -1 \\ -2 \end{bmatrix}, \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ u_1 \end{bmatrix}.$$

Below are the steps of Lemke's algorithm applied to this problem.

Phase I

Step 1: According to Phase I of Lemke's Algorithm, here we add the artificial variable x_0 that satisfy the following relationship, $s = Mx + ex_0 + q$, so the initial table is as follows.

	x_1	x_2	x_3	x_0	1
$s_1 =$	1	-1	-1	1	4
$s_2 =$	-1	1	-1	1	-1
$s_3 =$	1	1	0	1	-2

We make this table feasible by carrying out a Jordan elimination on the x_0 column (pivot column, s=4) and the row corresponding to the most negative entry in the last column (pivot row, r=3). Here $s = 4$ and $r = 3$. Since $B_{rs} = \frac{1}{A_{rs}}$ and $B_{rj} = \frac{-A_{rj}}{A_{rs}}$, $B_{is} = \frac{A_{is}}{A_{rs}}$ and $B_{rj} = A_{ij} - B_{is}A_{rj}$ we find the entries of the second table below.

	x_1	x_2	x_3	s_3	1
$s_1 =$	0	-2	-1	1	6
$s_2 =$	-2	0	-1	1	1
$x_0 =$	-1	-1	0	1	2

This table yields almost complementary solution $x_0 = 2$, $x_1 = 0$, $x_2 = 0$, $x_3 = 0$ and $s_1 = 0$, $s_2 = 1$, $s_3 = 0$.

Phase II

Step 2: In Phase I we obtained the following table.

	x_1	x_2	x_3	s_3	1
$s_1 =$	0	-2	-1	1	6
$s_2 =$	-2	0	-1	1	1
$x_0 =$	-1	-1	0	1	2

Since s_3 became non basic at the last pivot, here we choose x_3 as pivot column.

Minimum ratio test gives $min\left\{\frac{-6}{-1} = 6, \frac{-1}{-1} = 1\right\} = 1$.

Thus pivot row is $r = 2$ (from minimum ratio test). When $s = 3$ and $r = 2$ we find the entries in the third table by using the Jordan elimination. Using formulas indicated in Step 1 we obtain the following table:

	x_1	x_2	s_2	s_3	1
$s_1 =$	2	-2	1	0	5
$x_3 =$	-2	0	-1	1	1
$x_0 =$	-1	-1	0	1	2

This table yields almost complementary solution $x_0 = 2$, $x_1 = 0$, $x_2 = 0$, $x_3 = 1$ and $s_1 = 5$, $s_2 = 0$, $s_3 = 0$.

Step 3: By continuing the same process as in Step 2 we get $s = 2$ and $r = 3$. After performing Jordan elimination we obtain the following table.

	x_1	x_0	s_2	s_3	1
$s_1 =$	4	2	1	-2	1
$x_3 =$	-2	0	-1	1	1
$x_2 =$	-1	-1	0	1	2

This is a final table, because it contains a solution that is fully complementary, $x_0 = 0$, $x_1 = 0$, $x_2 = 2$, $x_3 = 1$ and $s_1 = 1$, $s_2 = 0$, $s_3 = 0$. Hence, the solution of the original problem (3.3) is $x_1 = 0$ and $x_2 = 2$.

Alternate Lemke Method We shall now proceed to show how Lemke's Method can be performed by avoiding traditional Jordan Exchanges on individual components but entire rows and columns are updated simultaneously.

Let us consider once again, example (3.3). We know that the initial table with artificial variable x^0 included is

	x_1	x_2	x_3	x_0	1
$s_1 =$	1	-1	-1	1	4
$s_2 =$	-1	1	-1	1	-1
$s_3 =$	1	1	0	1	-2

Now to make this table feasible:

(1) Pivot row (old) chosen by the most negative entry of the last column.

(2) Pivot element is the element of the pivot row (old) corresponding to the artificial variable, also identifying the pivot column.

Calculations: (Pivot, Row, and Column updates)

(3) Pivot $\text{Element}_{new} = \frac{1}{PivotElement_{old}}$

(4) Pivot $\text{Row}_{new} = \frac{-(PivotRow_{old})}{PivotElement_{old}}$ excluding elements corresponding to pivot column

(5) Pivot $\text{Column}_{new} = \frac{(PivotColumn_{old})}{PivotElement_{old}}$

(6) $\text{Row}_{new} = Row_{old} - (CorrespondingPivotElement_{new}) \times (PivotRow_{old})$ excluding element corresponding to pivot column

(7) Input pivot column_{new} where excluded column is located.

Step 6 updates all remaining elements of the table, and below we apply the above calculations to the given example.

(1)$PivotR\ ow_{old} = row3 = [1\ 1\ 0\ 1\ \ -2]$

(2)$Pivot\ Element_{old}$ =1, and pivot $\text{column}_{old} = [1\ 1\ 1]^T$

(3) Pivot $\text{Element}_{new} = \frac{1}{1} = 1$

(4)$Pivot\ Row_{new} = \frac{-[1\ 1\ 0\ *\ -2]}{1} = [-1\ \ -1\ 0\ \ *\ \ 2] = row3_{new}$

(5) Pivot $\text{Column}_{new} = \frac{[1\ 1\ 1]^T}{1}$

(6) Row_{1new} = [1 -1 -1 * 4]-(1)×[1 1 0 * -2] = [0 -2 -1 * 6]

Row_{2new} = [-1 1 -1 * -1]-(1)×[1 1 0 * -2] = [-2 0 -1 * 1]

Once the Pivot Column_{new} is input, the table is now feasible, and we have

	x_1	x_2	x_3	s_3	1
$s_1 =$	0	-2	-1	1	6
$s_2 =$	-2	0	-1	1	1
$x_0 =$	-1	-1	0	1	2

which satisfies the almost complementary solution specified by Phase I of Lemke's Method. Thus we can proceed to Phase 2 where all rules of Phase 2 hold and once again update the table by following steps 1 through 7 of (3.3). We can easily see that the same tables will be obtained. Repeat Lemke's phase 2 and (3.3) until a complementary solution or unbounded ray is determined.

CHAPTER 4

INFEASIBLE FULL NEWTON-STEP INTERIOR-POINT METHOD

The purpose of this chapter is to discuss and explain the IPM method for solving a monotone LCP. Step size calculations are not considered because this method utilizes full-Newton-steps. In this chapter we introduce and explain the concept of the IPM with full-Newton-steps, followed by the analysis of the convergence.

4.1 Main Idea of the Method

We consider the monotone LCP:

$$\begin{aligned} s &= Mx + q \\ x &\geq 0, s \geq 0 \\ x^T s = 0 &\Leftrightarrow xs = 0 \end{aligned} \tag{4.1}$$

where $M \in \mathbb{R}^{nxn}$ is a positive semidefinite matrix, $q \in \mathbb{R}^n$ is a vector. We say that (x, s) is an *ϵ-approximate solution* of (4.1) if

$$||s - Mx - q|| \leq \epsilon \;\; and \;\; x^T s \leq \epsilon \tag{4.2}$$

It is easy to see but important to note that because $x \geq 0, s \geq 0$

$$x^T s = 0 \Leftrightarrow xs = 0$$

where $xs = (x_1 s_1, ..., x_n s_n)$ represents the component-wise (Hadamard) product of vectors x, s. The main idea of IPM is to solve the system (4.1) using Newton's method. However, it is well known that Newton's method can "get stuck" at the *complementarity equation* $x^T s = 0$. In order for us to avoid this, we perturb the complementarity equation and consider the system:

$$\begin{aligned} s &= Mx + q \\ xs &= \mu e \end{aligned} \tag{4.3}$$

for some parameter $\mu \geq 0$. Another known fact is that for a positive definite matrix M, (4.3) has a unique solution $(x(\mu), s(\mu))$ for each $\mu > 0$. These (parametric) solutions are called μ-centers and the set of all μ-centers of (4.1) is called the *central path*. We can clearly see that when $\mu = 0$, we have found the solution for (4.1). Now, the main idea of IPM is to trace the central path by gradually reducing μ to 0. However, tracing the central path exactly is very inefficient; it is enough to trace it approximately, as long as the iterates are "not too far" from μ-centers. A clear understanding of "not too far" will be discussed more precisely later. The above outline of the IPM implicitly requires the existence and knowledge of a strictly feasible starting point (x^0, s^0), that is $s^0 = Mx^0 + q$ where $x^0 > 0, s^0 > 0$. The existence of a strictly feasible point is often called the *Interior-point condition* (IPC). However, finding a strictly feasible interior-point may be as difficult as solving the entire problem. Therefore, our goal in this thesis is to design an *Infeasible IPM*, this is an IPM that can start from an infeasible point and still converge and moreover converge relatively fast.

Let us consider an arbitrary starting point $(x^0, s^0) > 0$ such that $x^0 s^0 = \mu^0 e$ for some $\mu^0 > 0$. Most likely for this point, $s^0 \neq Mx^0 + q$ so we denote the residual as:

$$s^0 - Mx^0 - q = r^0. \tag{4.4}$$

The main idea of the *infeasible IPM* is to consider the corresponding perturbed LCP_ν

$$\begin{gathered} s - Mx - q = \nu r^0 \\ x \geq 0, \quad s \geq 0, \quad xs = 0 \end{gathered} \tag{4.5}$$

for any $0 < \nu \leq 1$. Note that when $\nu = 1$, (x^0, s^0) is a strictly feasible solution of $LCP_{\nu=1}$. Thus $LCP_{\nu=1}$ satisfies the IPC. The following lemma connects feasibility of the original LCP with feasibility of the corresponding perturbed LCP_ν.

Lemma 4.1.1. *If the original problem (4.1) is feasible, then the perturbed problem (4.5) is strictly feasible for* $\nu \in (0, 1]$

Proof. Suppose that (4.1) is feasible. Let $(\bar{x}, \bar{s})$ be a feasible solution, i.e.,

$$\bar{s} = M\bar{x} + q, \quad \bar{x} \geq 0, \quad \bar{s} \geq 0.$$

For $\nu \in (0, 1]$ consider convex combinations

$$x = (1 - \nu)\bar{x} + \nu x^0, \qquad s = (1 - \nu)\bar{s} + \nu s^0.$$

Note that $x, s > 0$ because $\bar{x}, \bar{s} \geq 0$ and $x^0, s^0 > 0$ and $\nu > 0$. We have

$$\begin{aligned}
s - Mx - q &= M((1-\nu)\bar{x} + \nu x^0) + q - ((1-\nu)\bar{s} + \nu s^0) \\
&= (1-\nu)M\bar{x} + \nu M x^0 + q - (1-\nu)\bar{s} - \nu s^0 \\
&= (1-\nu)(M\bar{x} - \bar{s}) + \nu(Mx^0 - s^0) + q \\
&= (1-\nu)(M\bar{x} + q - \bar{s} - q) + \nu(Mx^0 + q - s^0 - q) + q \\
&= (1-\nu)(-q) + \nu(r^0 - q) + q \\
&= (1-\nu)(-q) + \nu(-q) + \nu(r^0) + q \\
&= (-q)(1 - \nu + \nu) + \nu r^0 + q \\
&= -q + q + \nu r^0 \\
&= \nu r^0.
\end{aligned}$$

Thus (x, s) is strictly feasible for (4.5). $\square$

It is worth mentioning that $LCP_\nu \to LCP$ as $\nu \to 0$. Similarly, as for LCP (4.1), the perturbed problem LCP$_\nu$ (4.5) can be solved using IPM which would require solving the system

$$\begin{aligned} s - Mx - q &= \nu r^0 \\ xs &= \mu^+ e \end{aligned}$$

using Newton method. If M is p.s.d. this system has unique solutions $(x(\mu,\nu), s(\mu,\nu))$ of μ-centers for LCP_ν, which are also called *(μ,ν)-centers*. As before they form a central path for LCP_ν.

We seek not to find an exact solution of LCP_ν yet our goal is to locate an ϵ-approximate solution of the original problem LCP. That is achieved by finding an approximate solution (x,s) "close" to the (μ,ν)-center for a certain μ. Next, we simultaneously reduce the values of μ and ν for a certain parameter $\theta \in [0,1]$, called the *barrier parameter*, i.e.,

$$\begin{aligned} \mu^+ &= (1-\theta)\mu \\ \nu^+ &= (1-\theta)\nu \end{aligned}$$

As $\nu \to 0$ and $\mu \to 0$ we will obtain an ϵ-approximate solution of the original LCP. Since the initial μ is $\mu = \mu_0$ and the initial ν is $\nu = 1$, μ and ν are connected as follows:

$$\nu = \tfrac{\mu}{\mu_0}. \tag{4.6}$$

The *variance vector* defines the closeness of (x,s) to the μ-center and is denoted as follows, $v = \sqrt{\frac{xs}{\mu}}$. One can easily see that if (x,s) is a μ-center, which means that $xs = \mu e$, it immediately follows that $v = e$. Now, we define closeness (x,s) to the μ-center as $\delta(x,s;\mu) = \delta(v) = \frac{1}{2}||v - v^{-1}||$, here we notice that

$$\delta(v) = 0 \Leftrightarrow v = e \Leftrightarrow (x,s) \text{ is a } \mu\text{-center.}$$

As seen in figure 4.1 to obtain an approximate solution (x^+, s^+) in LCP$_{\nu^+}$ that is close to the μ^+-center, we perform one feasibility step and a few centering steps

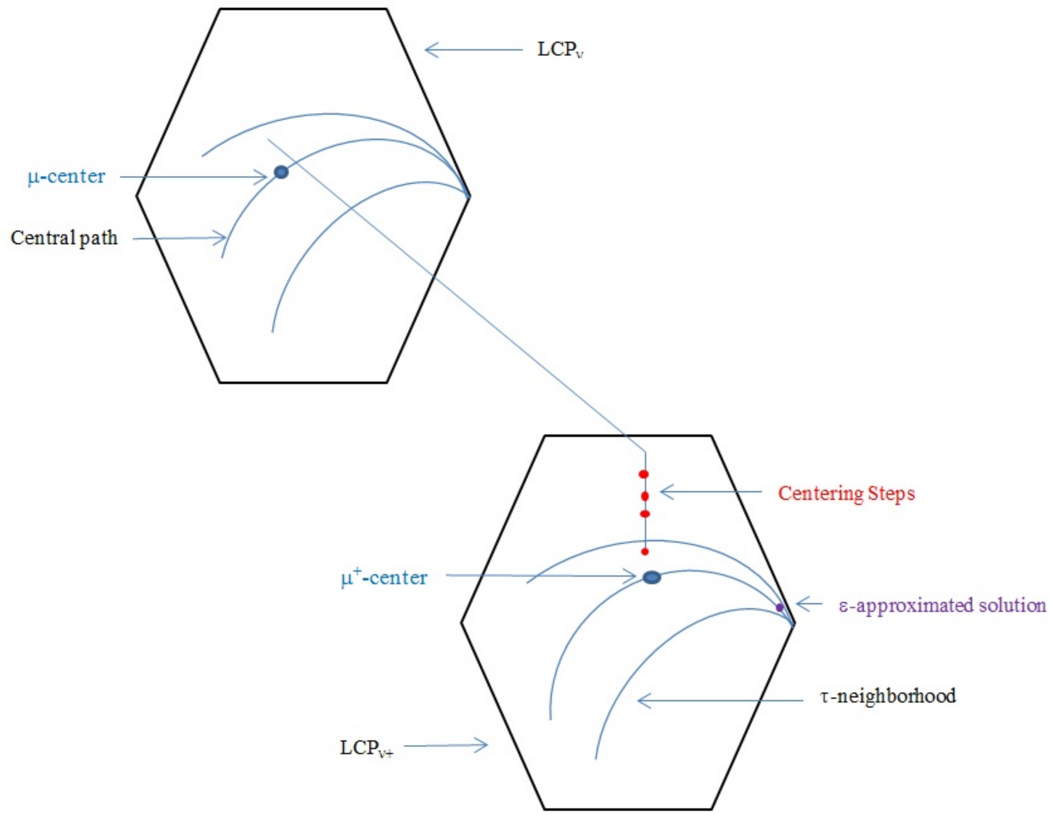

Figure 4.1: IIPM Graphical Representation

starting from the approximate solution (x, s) in LCP$_\nu$ close to μ-center. The feasibility step will assure that we obtained a solution (x^f, s^f) that is strictly feasible for LCP$_{\nu+}$ but may not be sufficiently close to the μ^+-center. Therefore, a single iteration consists of one feasibility step followed by several centering steps. In what follows we first describe the details of the feasibility step.

4.2 Feasibility Step

Let x, s be an approximate solution of LCP_{ν^+} that is known. Our goal is to find a strictly feasible solution of LCP_{ν^+}. Thus we want to approximately solve LCP_{ν^+} (4.5) using one iteration of Newton Method that will find the search direction $\Delta^f x, \Delta^f s$. We re-write (4.5) in the form:

$$F(x,s) = \begin{bmatrix} Mx + q - s - \nu r^0 \\ xs - \mu^+ e \end{bmatrix} = 0 \tag{4.7}$$

When we apply Newton method to (4.7) it yields:

$$\nabla F \begin{bmatrix} \Delta^f x \\ \Delta^f s \end{bmatrix} = -F(x,s) \tag{4.8}$$

where F is the Jacobian of F. The above system is equivalent to the following system:

$$\begin{aligned} \mathrm{M}\Delta^f x - \Delta^f s &= \theta\nu r^0, \\ s\Delta^f x + x\Delta^f s &= (1-\theta)\mu e - xs. \end{aligned} \tag{4.9}$$

Once the Newton directions $\Delta^f x, \Delta^f s$ are known, the feasible solution is obtained by performing a full-Newton update, i.e,

$$x^f = x + \Delta^f x$$

$$s^f = s + \Delta^f s$$

We want to be assured that x^f, s^f is strictly feasible and moreover, $\delta(x^f, s^f; \mu^+) < \frac{1}{\sqrt{2}}$.

4.3 Centering Step

Once (x^f, s^f) is obtained, we seek to obtain a solution (x, s) that is closer to μ^+-center than $\frac{1}{\sqrt{2}}$, that is we want to find (x, s) such that $\delta(x^f, s^f; \mu^+) \leq \tau$ for some "small" $\tau > 0$. This is achieved by performing a few centering steps within LCP_{ν^+} without moving to *new* LCP_{ν^+} denoted as $\text{LCP}_{\nu^{++}}$. Since we stay in LCP_{ν^+} we are not changing ν^+ nor μ^+ and therefore we can call ν^+ and μ^+ simply ν and μ. So they are named in this manner:

$$\nu \leftarrow \nu^+ \quad , \quad \mu \leftarrow \mu^+$$

Also, x^f, s^f can be called x, s and this is the starting point for our centering steps. Since x^f, s^f is strictly feasible we have

$$Mx^f + q - s^f = 0 \tag{4.10}$$

or because of renaming just

$$Mx + q - s = 0 \tag{4.11}$$

Once we have a strictly feasible solution, feasibility of the centering steps of the IPM is maintained. Thus one centering step consist of solving the system

$$\begin{aligned} Mx + q - s &= 0 \\ xs &= \mu e \end{aligned} \tag{4.12}$$

using Newton's method. The "centering" Newton direction is found by solving the following system

$$\begin{aligned} M\Delta^c x - \Delta^c s &= 0 \\ S\Delta^c x + X\Delta^c s &= \mu e - xs \end{aligned} \tag{4.13}$$

Then the new centering solution is obtained by taking a full Newton-step

$$\begin{aligned} x^c &= x + \Delta^c x \\ s^c &= s + \Delta^c s \end{aligned}$$

We will show that (x^c, s^c) is closer to μ-center than (x, s), actually we will show that proximity to the μ-center is reduced quadratically. This outline is summarized in the following algorithm.

Infeasible Full Newton-step Interior-Point Algorithm for LCP

Input:

Determine input parameters:

threshold parameter $\tau > 0$,

fixed barrier update parameter θ, $0 < \theta < 1$,

accuracy parameter $\epsilon > 0$.

begin

Set $\mu^0 = \zeta_P \zeta_D$, $\zeta_P > 0$ and $\zeta_D > 0$

$x^0 s^0 = \mu^0 e$

$\nu = 1$

while $max(x^T s,\ ||s - Mx - q||) \geq \epsilon$ **do**

Feasibility Step

Calculate direction $(\Delta^f x, \Delta^f s)$ by solving (4.9);

Update $x := x + \Delta x$ and $s := s + \Delta s$;

Update $v := \sqrt{\frac{xs}{\mu}}$;

Calculate: $\delta(v) = \frac{1}{2}||v - v^{-1}||$;

$\mu := (1 - \theta)\mu$;

$\nu := (1 - \theta)\nu$;

Centering Step

while $\delta(v) > \tau$ **do**

Calculate original direction $(\Delta x, \Delta s)$ by solving (4.13);

Update $x := x + \Delta x$ and $s := s + \Delta s$;

Update $v := \sqrt{\frac{xs}{\mu}}$;

end do

end do

end

Table 4.1: Infeasible Full Newton-step Interior-Point Algorithm for LCP

CHAPTER 5

ANALYSIS OF THE ALGORITHM

This purpose of this chapter is to analyze convergence and to estimate the number of iterations needed to find an ϵ-approximate solution of LCP. We start out with the analysis of the feasibility step, followed by the analysis of the centering steps, and conclude with overall number of iterations.

5.1 Feasibility Step

Let us recall from Chapter 4 the system (4.9) for the feasible step.

$$\begin{aligned} \mathrm{M}\Delta^f x - \Delta^f s &= \theta\nu r^0, \\ s\Delta^f x + x\Delta^f s &= (1-\theta)\mu e - xs. \end{aligned} \tag{5.1}$$

In order to analyze the above system it is useful to transform it into an equivalent system using the following scaled directions.

$$v := \sqrt{\frac{xs}{\mu}}, \quad d_x := \frac{v\Delta x}{x}, \quad d_s := \frac{v\Delta s}{s} \tag{5.2}$$

Note that the pair (x, s) coincides with the μ-center $(x(\mu), s(\mu))$ if and only if $v = e$. Substitution of (5.2) into the above system yields

$$\begin{aligned} M\tfrac{xd_x}{v} - \tfrac{sd_s}{v} &= \theta\nu r^0 \\ S\tfrac{xd_x}{v} + X\tfrac{sd_s}{v} &= (1-\theta)\mu e - xs \end{aligned} \tag{5.3}$$

Writing system (5.3) in matrix form we have:

$$MV^{-1}Xdx - SV^{-1}ds = \theta\nu r^0 \tag{5.4}$$

$$SV^{-1}Xdx + XV^{-1}ds = (1-\theta)\mu e - xs, \tag{5.5}$$

where

$$X = diag(x), \quad V^{-1} = diag(v^{-1}), \quad D_x = diag(d_x)$$

$$S = diag(s), \quad V = diag(v), \quad D_s = diag(d_s).$$

Left multiplication of (5.4) by $S^{-1}V$ yields:

$$\begin{aligned} S^{-1}V(MV^{-1}Xd_x) - S^{-1}V(d_s) &= (S^{-1}V)(\theta\nu r^0) \\ S^{-1}MXd_x - d_s &= \theta\nu S^{-1}Vr^0 \\ S^{-1/2}X^{1/2}MS^{-1/2}X^{1/2}d_x - ds &= \theta\nu S^{-1}X^{1/2}S^1\mu^{-1/2}r^0 \end{aligned} \tag{5.6}$$

Let $D := S^{-1/2}X^{1/2}$, therefore the third equation in (5.6) is written as

$$MDd_x - ds = \theta\nu D\mu^{-1/2}r^0. \tag{5.7}$$

Let $\widetilde{M} := DMD$ therefore (5.8) is written as

$$\widetilde{M}\mathrm{d}_x - ds = \theta\nu D\mu^{-1/2}r^0. \tag{5.8}$$

If we left multiply (5.5) by $S^{-1}X^{-1}V$ we obtain

$$\begin{aligned} S^{-1}X^{-1}V(SXV^{-1}d_x + XSV^{-1}d_s) &= S^{-1}X^{-1}V[(1-\theta)\mu e - xs] \\ d_x + d_s &= X^{-1}S^{-1}V[(l-\theta)\mu e - xs] \\ d_x + d_s &= \tfrac{\mu}{xs}\sqrt{\tfrac{xs}{\mu}}(1-\theta)e - X^{-1}S^{-1}XSV \\ d_x + d_s &= \sqrt{\tfrac{\mu}{xs}}(1-\theta)e - v \\ d_x + d_s &= (1-\theta)v^{-1} - v. \end{aligned}$$

Thus, system (5.3) transforms into the following system

$$\begin{aligned} \widetilde{M}d_x - ds &= \theta\nu\tfrac{1}{\sqrt{\mu}}Dr^0 \\ d_x + d_s &= (1-\theta)v^{-1} - v \end{aligned} \tag{5.9}$$

Recall that after the feasibility step,

$$x^f = x + \Delta^f x, \quad s^f = s + \Delta^f s.$$

where $\Delta^f x,\ \Delta^f s$ are calculated from system (5.1). We want to guarantee that $x^f > 0$, $s^f > 0$, so our goal is to find the condition that will guarantee strict feasibility of x^f, s^f. Recall that we start with $(x, s; \mu)$ such that $\delta(x, s; \mu) < \tau$. We reduce μ to $\mu^+ = (1-\theta)\mu$. Now , using $v := \sqrt{\frac{xs}{\mu}}$, $\quad d_x^f := \frac{v\Delta^f x}{x}$, $\quad d_s^f := \frac{v\Delta^f s}{s}$ we obtain

$$\begin{aligned} x^f s^f &= xs + (x\Delta^f x + s\Delta^f s) + \Delta^f x \Delta^f s \\ &= xs + (1-\theta)\mu e - xs + \Delta^f x \Delta^f s \\ &= (1-\theta)\mu e + \Delta^f x \Delta^f s \\ &= (1-\theta)\mu e + \tfrac{xs}{v^2} d_x^f d_s^f \\ &= (1-\theta)\mu e + \tfrac{xs}{\frac{xs}{\mu}} d_x^f d_s^f \\ &= (1-\theta)\mu e + \mu d_x^f d_s^f, \end{aligned} \tag{5.10}$$

which implies

$$\begin{aligned} (v^f)^2 &= \tfrac{x^f s^f}{\mu^+} \\ &= \tfrac{(1-\theta)e + d_x^f d_s^f}{(1-\theta)} \tfrac{\mu}{\mu} \\ &= e + \tfrac{d_x^f d_s^f}{1-\theta}, \end{aligned} \tag{5.11}$$

or equivalently

$$(v^f)_i^2 = 1 + \tfrac{d_{x_i}^f d_{s_i}^f}{1-\theta}. \tag{5.12}$$

Lemma 5.1.1. *Iterates (x^f, s^f) are strictly feasible if and only if $(1-\theta)e + d_x^f d_s^f > 0$.*

Proof. $(:\Rightarrow)$ If $x^f$and s^f are both positive then $(1-\theta)e + d_x d_s > 0$

$(\Leftarrow:)$

let

$$x^0 = x, \quad s^0 = s$$

$$x^1 = x + \Delta^f x$$

$$s^1 = s + \Delta^f s$$

therefore, $x^0 s^0 > 0$. We need to show x^1 and s^1 are nonnegative if $x^\alpha s^\alpha$ is positive for all $\alpha \in (0, 1)$.

$$\begin{aligned}
x^\alpha s^\alpha &= (x + \alpha\Delta^f x)(s + \alpha\Delta^f s) \\
&= xs + x\alpha\Delta^f s + s\alpha\Delta^f x + \alpha^2\Delta^f x\Delta^f s \\
&= xs + \alpha(x\Delta^f x + s\Delta^f s) + \alpha^2\Delta^f x\Delta^f s \\
&= xs + \alpha((1-\theta)\mu e - xs) + \alpha^2\Delta^f x\Delta^f s \\
&= xs + \alpha(1-\theta)\mu e - \alpha xs + \alpha^2\mu d_x^f d_s^f \\
&= xs + \alpha\mu e - \alpha\theta\mu e - \alpha xs + \alpha^2\mu d_x^f d_s^f \\
&= xs(1-\alpha) + \alpha\mu e(1-\theta) + \alpha^2 d_x^f d_s^f \\
&= \mu[(1-\alpha)v^2 + \alpha e(1-\theta) + \alpha^2 d_x^f d_s^f]
\end{aligned}$$

Suppose $(1-\theta)e + d_x^f d_s^f > 0$ then $d_x^f d_s^f > -(1-\theta)e$. Substitution yields

$$\begin{aligned}
x^\alpha s^\alpha &> \mu[(1-\alpha)v^2 + \alpha e(1-\theta) + \alpha^2(1-\theta)e] \\
&= \mu[(1-\alpha)v^2 + \alpha e - \alpha e\theta - \alpha^2 e + \alpha^2\theta e] \\
&= \mu[(1-\alpha)v^2 + \alpha e(1-\alpha) - \alpha e\theta(1-\alpha)] \\
&= \mu(1-\alpha)[v^2 + \alpha e - \alpha e\theta] \\
&= \mu(1-\alpha)[v^2 + \alpha(1-\theta)e].
\end{aligned}$$

Since $v^2, e > 0$ this implies $x^\alpha s^\alpha > 0$ for $\alpha \in [0, 1)$. Therefore none of the entries of x^α and s^α vanish for $\alpha \in [0, 1)$. Since $x^0, s^0 > 0$, this implies that $x^\alpha > 0$ and $s^\alpha > 0$ for $\alpha \in [0, 1)$. So by continuity the vectors x^1 and s^1 cannot have negative entries. Assuming $(1-\theta)e + d_x d_s > 0$. Similarly $x^\alpha s^\alpha > \mu[(1-\alpha)v^2 + \alpha e(1-\theta) + \alpha^2(1-\theta)e]$ implies $x^1 s^1 > 0$; therefore by continuity x^1 and s^1 must be positive. $\square$

Lemma 5.1.2. (x^f, s^f) *are strictly feasible if* $||d_x^f d_s^f||_\infty < 1 - \theta$

Proof. By Lemma 5.1.1 (x^f, s^f) is strictly feasible if and only if

$$\begin{aligned} (1-\theta)e + d_x^f d_s^f &> 0 \\ (1-\theta) + d_{x_i}^f d_{s_i}^f &> 0 \\ d_{x_i}^f d_{s_i}^f &> -(1-\theta). \end{aligned}$$

Given the definiton of ∞-norm, $\left\| d_x^f d_s^f \right\|_\infty = max \left\{ \left| d_{x_i}^f d_{s_i}^f \right| : i = 1, ..., n \right\}$ the assumption $\left\| d_x^f d_s^f \right\|_\infty < 1-\theta$ can be written as $\left\| d_{x_i}^f d_{s_i}^f \right\| < 1-\theta$ or equivalently

$$-(1-\theta) < d_{x_i}^f d_{s_i}^f < (1-\theta).$$

The left inequality above can be written as $d_{x_i}^f d_{s_i}^f + (1-\theta) > 0$ for any i or
$d_x^f d_s^f (1-\theta)e > 0$.
By Lemma (5.1.1) this implies that (x^f, s^f) is strictly feasible. $\square$

Now we seek to find conditions that would lead to the required upper-bound for $\delta(x^f, s^f; \mu^+) \leq \frac{1}{\sqrt{2}}$.

Lemma 5.1.3. *If* $\left\| d_x^f d_s^f \right\|_\infty < 1-\theta$ *then* $4\delta^2(v^f) \leq \frac{\left\| d_x^f d_s^f \right\|^2}{1-\left\| d_x^f d_s^f \right\|_\infty}$.

Proof. Recall that

$$\begin{aligned} \delta(v^f) &= \delta(x^f, s^f, \mu^+) \\ &= \tfrac{1}{2} \left\| v^f - (v^f)^{-1} \right\| \end{aligned}$$

Then we have

$$\begin{aligned} 4\delta^2(v^f) &= 4(\tfrac{1}{4}||v^f - (v^f)^{-1}||^2) \\ &= ||v^f - (v^f)^{-1}||^2 \\ &= \textstyle\sum_{i=1}^n (v_i^f - \frac{1}{v_i^f})^2 \\ &= \textstyle\sum_{i=1}^n (v_i^f)^2 - 2 + (\frac{1}{v_i^f})^2 \\ &= \textstyle\sum_{i=1}^n (1 + \frac{d_{x_i}^f d_{s_i}^f}{1-\theta} + \frac{1}{1+\frac{d_{x_i}^f d_{s_i}^f}{1-\theta}} - 2), \end{aligned}$$

where the last equality is obtained using (5.12).

Let us denote $z_i := \frac{d^f_{x_i} d^f_{s_i}}{1-\theta}$; therefore, we have

$$\begin{aligned}
\textstyle\sum_{i=1}^{n}(1 + z_i + \frac{1}{1+z_i} - 2) &= \textstyle\sum_{i=1}^{n}(z_i + \frac{1}{1+z_i} - 1) \\
&= \textstyle\sum_{i=1}^{n}(\frac{z_i(1+z_i)+1}{1+z_i} - 1) \\
&= \textstyle\sum_{i=1}^{n}(\frac{z_i^2+z_i+1}{1+z_i} - 1) \\
&= \textstyle\sum_{i=1}^{n}(\frac{z_i^2}{1+z_i})
\end{aligned} \tag{5.13}$$

Now $z_i \leq |z_i|$ and $z_i \geq -|z_i|$. So $|z_i| \leq \|z_i\|_\infty$ this implies

$$\begin{aligned}
1 - |z_i| &\geq 1 - \|z_i\|_\infty \\
1 + z_i &\geq 1 - |z_i| .
\end{aligned}$$

Therefore,

$$\frac{1}{1+z_i} \geq \frac{1}{1-z_i} \geq \frac{1}{1+z_{i\infty}}. \tag{5.14}$$

By substituting (5.14) into (5.13) we get

$$\begin{aligned}
\delta^2(v^f) &\leq \textstyle\sum_{i=1}^{n}(\frac{z_i^2}{1-|z_i|}) \\
&\leq \textstyle\sum_{i=1}^{n}(\frac{z_i^2}{1-|z_i|_\infty}) \\
&= \frac{1}{1-\|z_i\|_\infty}(\|z\|_2^2) \\
&= \frac{(\|z\|_2^2)}{1-\|z_i\|_\infty} \\
&= \frac{\left\|\frac{d^f_{x_i} d^f_{s_i}}{1-\theta}\right\|_2^2}{1-\left\|\frac{d^f_{x_i} d^f_{s_i}}{1-\theta}\right\|_\infty}
\end{aligned} \tag{5.15}$$

$\square$

Norm Facts

We introduce the following known facts about the norms to assist us in the analysis of the feasibility step.

$$\begin{aligned} \left\| d_x^f d_s^f \right\|_\infty &\le \left\| d_x^f d_s^f \right\| \\ &\le \left\| d_x^f \right\| \left\| d_s^f \right\| \\ &\le \tfrac{1}{2}(\left\| d_x^f \right\|^2 + \left\| d_s^f \right\|^2) \end{aligned} \tag{5.16}$$

Where $\|\|$ represents the 2-norm. Using (5.16), the last equation of (5.15) can be written as

$$4\delta^2(v^f) \le \frac{\frac{1}{4}\left(\frac{\left\| d_x^f \right\|^2 + \left\| d_s^f \right\|^2}{1-\theta}\right)^2}{1-\frac{1}{2}\frac{\left\| d_x^f \right\|^2 + \left\| d_s^f \right\|^2}{1-\theta}}$$

Recall that we want $\delta(v^f) \le \frac{1}{\sqrt{2}}$ which implies $4\delta^2(v^f) \le 2$. Using Lemma 5.1.3 this will be satisfied if

$$\frac{\frac{1}{4}\left(\frac{\left\| d_x^f \right\|^2 + \left\| d_s^f \right\|^2}{1-\theta}\right)^2}{1-\frac{1}{2}\frac{\left\| d_x^f \right\|^2 + \left\| d_s^f \right\|^2}{1-\theta}} \le 2. \tag{5.17}$$

This implies that

$$\tfrac{1}{2}\tfrac{||d_x^f||^2 + ||d_s^f||^2}{1-\theta} \le 1 \tag{5.18}$$

However, we should not forget that Lemma 5.1.3 must hold, and therefore (5.18) holds if $\left\| \frac{d_x d_s}{1-\theta} \right\|_\infty < 1$. So the problem of finding conditions for $\delta(v^f) \le \frac{1}{\sqrt{2}}$ to hold reduces to finding an upper bound on

$$\left\| d_x^f \right\|^2 + \left\| d_s^f \right\|^2. \tag{5.19}$$

To assist us in finding this bound, we have the following condition: $\left\| d_x^f d_s^f \right\|_\infty < 1 - \theta$

$$\begin{aligned} &\Rightarrow \tfrac{\left\| d_x^f d_s^f \right\|}{1-\theta} < 1 \\ &\Rightarrow \tfrac{1}{2}\tfrac{\left\| d_x^f \right\|^2 + \left\| d_s^f \right\|^2}{1-\theta} \le 1 \end{aligned}$$

The question we ask ourselves now is "how do we find the upper bound for $\left\| d_x^f \right\|^2 + \left\| d_s^f \right\|^2$?" In order to do so we need the following lemma.

Lemma 5.1.4. *Given a system*

$$\widetilde{M}u - z = \tilde{a}, \quad u + z = \tilde{b} \tag{5.20}$$

the following hold

$$\begin{array}{ll} (1) & Du = (1 + DMD)^{-1}(a + b), \quad Dz = (b - Du) \\ (2) & \|Du\| \leq \|a + b\| \\ (3) & \|Du\|^2 + \|Dz\|^2 \leq \|b\|^2 + 2\|a + b\|\|a\|. \end{array}$$

where D, b, a, and $\widetilde{M}$ are defined as follows:

$$D := S^{-1/2}X^{1/2}, \quad b := D\tilde{b}, \quad a := D\tilde{a} \quad \textit{and} \quad \widetilde{M} := DMD.$$

Proof. Left multiply both equations in (5.20) by D which gives us

$$\begin{aligned} DMDDu - Dz &= a \\ Du + DZ &= b, \end{aligned} \tag{5.21}$$

and by adding the 2 above equations we deduce equation (1) of Lemma 5.1.4. Since the matrix $I + DMD$ is positive definite, inequality (2) of Lemma 5.1.4 follows. From (5.21) and by using Cauchy-Schwartz inequality and inequality (2) of Lemma 5.1.4 and the positive semidefiniteness of DMD we have

$$\begin{aligned} \|Du\|^2 + \|Dz\|^2 &= \|Du + Dz\|^2 - 2(Du)^T Dz \\ &= \|b\|^2 - 2(Du)^T(DMDDu - a) \\ &= \|b\|^2 - 2(Du)^T DMDDu + 2(Du)^T a \\ &\leq \|b\|^2 + 2\|Du\|\|a\| \\ &\leq \|b\|^2 + 2\|a + b\|\|a\| \end{aligned}$$

□

We will apply the above lemma to the system (5.9)

$$\begin{aligned} \widetilde{M}d_x^f - d_s^f &= \theta\nu D r^0 \tfrac{1}{\sqrt{\mu}} \\ d_x^f + d_s^f &= (1-\theta)v^{-1} - v. \end{aligned}$$

Let

$$\begin{aligned} \tilde{a} : &= D(\theta\nu D r^0 \tfrac{1}{\mu}) = D^2(\theta\nu r^0 \tfrac{1}{\mu}) \\ \tilde{b} : &= D((1-\theta)v^{-1} - v) \\ u : &= d_x^f \\ z : &= d_s^f. \end{aligned}$$

Then system (5.9) transforms to system (5.20). Substituting into equation (3) of Lemma 5.1.4 we have

$$\begin{aligned} \left\|Dd_x^f\right\|^2 + \left\|Dd_s^f\right\| &\leq \|D[(1-\theta)v^{-1} - v]\|^2 \\ &+ 2\left\|D^2(\theta\nu r^0 \tfrac{1}{\sqrt{\mu}} + D((1-\theta)v^{-1} - v)\right\| \left\|D^2\theta\nu r^0 \tfrac{1}{\sqrt{\mu}}\right\| \end{aligned} \tag{5.22}$$

Using norm properties, we have that

$$\begin{aligned} \left\|Dd_x^f\right\| \leq \|D\| \left\|d_x^f\right\| \quad &, \quad \left\|Dd_s^f\right\| \leq \|D\| \left\|d_s^f\right\| \\ \left\|D^2\nu r^0 \tfrac{1}{\sqrt{\mu}}\right\| &\leq \|D\|^2 \left\|\theta\nu r^0\right\| \tfrac{1}{\sqrt{\mu}} \\ \|D[(1-\theta)v^{-1} - v]\| &\leq \|D\| \left\|(\theta\nu r^0 \tfrac{1}{\sqrt{\mu}}\right\| \end{aligned} \tag{5.23}$$

where $\|D\|$ represents a matrix norm. Thus

$$\left\|Dd_x^f\right\|^2 + \left\|Dd_s^f\right\|^2 \leq \|D\|^2 \left(\left\||d_x^f\right\|^2 + \left\|d_s^f\right\|^2\right) \tag{5.24}$$

Using inequalities (5.23) and (5.24), inequality (5.22) becomes

$$\begin{aligned} \|D\|^2 \left(\left\||d_x^f\right\|^2 + \left\|d_s^f\right\|^2\right) &\leq \|D\|^2 \left\|[(1-\theta)v^{-1} - v\right\|^2 \\ &+ 2\left(\left\|D\theta\nu r^0 \tfrac{1}{\sqrt{\mu}}\right\| + \|(1-\theta)v^{-1} - v\|\right) \left\|D\theta\nu r^0 \tfrac{1}{\sqrt{\mu}}\right\| \end{aligned} \tag{5.25}$$

Cancelling $\|D\|^2$ we get

$$\begin{aligned} \left\| |d_x^f \right\|^2 + \left\| d_s^f \right\|^2 &\leq \|(1-\theta)v^{-1} - v\|^2 \\ &+ 2\left(\left\| D\theta\nu r^0 \tfrac{1}{\sqrt{\mu}} \right\| + \|(1-\theta)v^{-1} - v\|\right) \left\| D\theta\nu r^0 \tfrac{1}{\sqrt{\mu}} \right\|. \end{aligned} \tag{5.26}$$

We now seek an upper bound for

$$\left\| \theta\nu D r^0 \right\| \tag{5.27}$$

and

$$\left\| (1-\theta)v^{-1} - v \right\|. \tag{5.28}$$

First we give a bound for (5.27). We have

$$\begin{aligned} \left\| \tfrac{\theta\nu}{\sqrt{\mu}} D r^0 \right\| &= \tfrac{\theta\nu}{\sqrt{\mu}} \|Dr^0\| \\ &= \tfrac{\theta\nu}{\sqrt{\mu}} \left\| X^{1/2} S^{-1/2} r^0 \right\| \\ &= \tfrac{\theta\nu}{\sqrt{\mu}} \left\| \sqrt{\tfrac{x}{s}} r^0 \right\|, \qquad \to \nu = \tfrac{\mu}{\mu_0} \\ &\leq \tfrac{\theta}{\sqrt{\mu}} \tfrac{\mu}{\mu_0} \left\| \sqrt{\tfrac{x}{s}} r^0 \right\|_1 \\ &= \tfrac{\theta}{\mu_0} \sqrt{\mu} \left\| \sqrt{\tfrac{x}{s}} r^0 \right\|_1 \\ &= \tfrac{\theta}{\mu_0} \left\| \sqrt{\tfrac{\mu x}{s}} r^0 \right\|_1 \\ &= \tfrac{\theta}{\mu_0} \left\| \sqrt{\tfrac{\mu}{xs}} x r^0 \right\|_1 \qquad \to \left| \tfrac{1}{v_i} x_i r_i^0 \right| \leq \tfrac{1}{v_{min}} |x_i r_i^0| \leq \tfrac{1}{v_{min}} |r_i^0| \, |x_i^0| \\ &= \tfrac{\theta}{\mu_0} \tfrac{1}{v_{min}} \|x r^0\|_1 \end{aligned}$$

therefore

$$\left\| \tfrac{\theta\nu}{\sqrt{\mu}} D r^0 \right\| \leq \tfrac{\theta}{\mu_0} \tfrac{1}{v_{min}} \|r^0\|_\infty \|x\|_1 . \tag{5.29}$$

Since we assumed $x^0 = \zeta_p e$ and $s^0 = \zeta_d e$ we have that $\mu_0 = \zeta_p \zeta_d$. We can choose ζ_p and ζ_d such that $\|x^0\|_\infty \leq \zeta_p$ and $\|s^0\|_\infty \leq \zeta_d$. Then we have

$$\begin{aligned}
r^0 &= s^0 - Mx^0 - q \\
&= \zeta_d e - \zeta_p Me - q \\
&= \zeta_d \left(e - \tfrac{\zeta_p}{\zeta_d} Me - \tfrac{1}{\zeta_d} q \right) \\
\|r^0\|_\infty &= \zeta_d \left\| e - \tfrac{\zeta_p}{\zeta_d} Me - \tfrac{1}{\zeta_d} q \right\|_\infty \\
&\leq \zeta_d \left(1 + \tfrac{1}{\zeta_d} \zeta_p \|Me\|_\infty + \tfrac{1}{\zeta_d} \|q\|_\infty \right).
\end{aligned}$$

By assuming $max \{ \|s^\alpha\|_\infty, \|Me\|_\infty, \|q\|_\infty \} \leq \zeta_d$ the last inequality above becomes

$$\|r^0\|_\infty \leq \zeta_d (1 + 1 + 1) = 3\zeta_d.$$

Thus, (5.29) becomes

$$\begin{aligned}
\left\| \tfrac{\theta \nu}{\sqrt{\mu}} D r^0 \right\| &\leq \tfrac{\theta}{\mu_0} \tfrac{1}{v_{min}} 3\zeta_d \|x\|_1 \\
= \tfrac{\theta}{\zeta_p \zeta_d} \tfrac{1}{v_{min}} 3\zeta_d \|x\|_1 &= \tfrac{3\theta}{\zeta_p} \tfrac{\|x\|_1}{v_{min}}.
\end{aligned} \tag{5.30}$$

Now, we give an upper bound for (5.28). We have

$$\begin{aligned}
\|(1-\theta) v^{-1} - v\|^2 &= \|(1-\theta) v^{-1}\|^2 - 2(1-\theta)(v^{-1})^T(v) + \|v\|^2 \\
&= (1-\theta)^2 \|v^{-1}\|^2 - 2(1-\theta)n + \|v\|^2 \\
&= (1-\theta)^2 \|v^{-1}\|^2 - 2n + \|v\|^2 + 2n\theta \\
&\leq \|v^{-1}\|^2 - 2n + \|v\|^2 + 2\theta n \\
&= \|v^{-1} - v\|^2 + 2\theta n \\
&= 4\delta^2(v) + 2\theta n
\end{aligned} \tag{5.31}$$

Thus

$$\|(1-\theta) v^{-1} - v\| = \sqrt{4\delta^2(v) + 2\theta n}.$$

By substituting equations (5.30) and (5.31) into (5.25) we get

$$\left\|d_x^f\right\|^2 + \left\|d_s^f\right\|^2 \leq (4\delta^2(v) + 2\theta n) + 2\left(\frac{3\theta}{\zeta_p}\frac{\|x\|_1}{v_{min}} + \sqrt{4\delta^2(v) + 2\theta n}\right)\frac{3\theta}{\zeta_p}\frac{\|x\|_1}{v_{min}}. \qquad (5.32)$$

Next, we need an upper bound on $\|x\|_1$ and a lower bound on v_{min}. This is given in the lemma below.

Lemma 5.1.5.

$$(1) \quad q^{-1}(\delta) \leq v_i \leq q(\delta)$$

$$(2) \quad \|x\|_1 \leq (2 + q(\delta))n\zeta_p, \quad \|s\|_1 \leq (2 + q(\delta))n\zeta_p$$

where

$$q(\delta) = \delta + \sqrt{\delta^2 + 1}.$$

Proof. Since v_i is positive for each i, we have

$$-2\delta v_i \leq 1 - v_i^2 \leq 2\delta v_i$$

This implies

$$v_i^2 - 2\delta v_i - 1 \leq 0 \leq v_i^2 + 2\delta v_i - 1$$

Rewriting this, we have

$$(v_i - \delta)^2 - 1 - \delta^2 \leq (v_i + \delta)^2 - 1 - \delta^2$$

we obtain

$$(v_i - \delta)^2 \leq 1 + \delta^2 \leq (v_i + \delta)^2,$$

which implies

$$v_i - \delta \leq |v_i - \delta| \leq \sqrt{1 + \delta^2} \leq v_i + \delta.$$

Thus we arrive at

$$-\delta + \sqrt{1 + \delta^2} \leq v_i \leq \delta + \sqrt{1 + \delta^2} = q(\delta).$$

For the expression on the left-hand side we write

$$-\delta + \sqrt{1+\delta^2} = \frac{1}{\delta+\sqrt{1+\delta^2}} = \frac{1}{q(\delta)}$$

thus proving number (1) of the above lemma. To prove (2), since x, s, x^* and s^* are positive, it implies that $s^T x^* + x^T s^*$ is positive. Therefore,

$$(s^0)^T x + (x^0)^T s \leq v(x^0)^T s^0 + \frac{x^T s}{v} + (1-v)((s^0)^T x^* + (x^0)^T s^*)$$

since $x^0 = \zeta_p e$, $s^0 = \zeta_D e$, $\|x^*\|_\infty \leq \zeta_p$ and $\|s^*\|_\infty \leq \zeta_D$, we have

$$(s^0)^T x^* + (x^0)^T s^* \leq \zeta_P(e^T s^0) + \zeta_D(e^T x^0) = 2n\zeta_P\zeta_D.$$

Also $(x^0)^T s^0 = n\zeta_P\zeta_D$. Hence we get

$$\begin{aligned}
(s^0)^T x + (x^0)^T s &\leq \frac{x^T s}{v} + 2n\zeta_p\zeta_D - vn\zeta_p\zeta_D \\
&\leq \frac{x^T s}{v} + 2n\zeta_p\zeta_D \\
&= \frac{\mu(e^T v^2)}{v} + 2n\zeta_p\zeta_D \\
&= \zeta_p\zeta_D(e^T v^2) + 2n\zeta_p\zeta_D,
\end{aligned}$$

where the last equality follows because of $v = \frac{\mu}{\mu^0}$ and $\mu^0 = \zeta_p\zeta_D$. Now,

$$\begin{aligned}
\zeta_p\zeta_D(e^T v^2) + 2n\zeta_p\zeta_D &= \zeta_p\zeta_D(e^T v^2 + 2n) \\
&= \zeta_p\zeta_D(\textstyle\sum v_i^2 + 2n) \\
&\leq \zeta_p\zeta_D(\textstyle\sum q^2(\delta) + 2n) \\
&= \zeta_p\zeta_D(q^2(\delta)\textstyle\sum 1 + 2n) \\
&= \zeta_p\zeta_D(q^2(\delta)n + 2n) \\
&= \zeta_p\zeta_D n(q^2(\delta) + 2).
\end{aligned}$$

Therefore, $(s^0)^T x + (x^0)^T s \leq (q(\delta)^2 + 2)n\zeta_p\zeta_D$. Since x^0, s^0, x and s are positive we obtain

$$\begin{aligned}
(s^0)^T x &\leq (q^2(\delta) + 2)n\zeta_p\zeta_D \\
(x^0)^T x &\leq (q^2(\delta) + 2)n\zeta_p\zeta_D.
\end{aligned}$$

Moreover, since $x^0 = \zeta_p e$ and $s^0 = \zeta_D e$, we obtain

$$\begin{aligned} \|x\|_1 &\leq (q^2(\delta)+2)n\zeta_p \\ \|s\|_1 &\leq (q^2(\delta)+2)n\zeta_p, \end{aligned}$$

thus part (2) is proven, which concludes the proof of the lemma. $\square$

Using Lemma 5.1.5, (5.32) becomes

$$\begin{aligned} &\left\|d_x^f\right\|^2 + \left\|d_s^f\right\|^2 \\ &\leq (4\delta^2 + 2\theta n) + \tfrac{18\theta^2}{\zeta_p^2}\tfrac{\|x\|_1^2}{v_{min}^2} + \tfrac{6\theta}{\zeta_p}\sqrt{4\delta^2+2\theta n}\tfrac{\|x\|_1}{v_{min}} \\ &= (4\delta^2 + 2\theta n) + \tfrac{18\theta^2}{\zeta_p^2}\Big((2+q(\delta))^2 n^2 \zeta_p^2)q^2(\delta) + \tfrac{6\theta}{\zeta_p}\sqrt{4\delta^2+2\theta n}(2+q(\delta))n\zeta_p q(\delta)\Big) \\ &= (4\delta^2 + 2\theta n) + 18\theta^2 n^2 (2+q(\delta))^2 q^2(\delta) + 6\theta n\sqrt{4\delta^2+2\theta n}(2+q(\delta))q(\delta) \end{aligned}$$

Therefore,

$$\begin{aligned} &\left\|d_x^f\right\|^2 + \left\|d_s^f\right\|^2 \\ &\leq (4\delta^2 + 2\theta n) + (18\theta^2 n^2 (2+q(\delta))^2 q^2(\delta) + 6\theta n\sqrt{4\delta^2+2\theta n}(2+q(\delta))q(\delta). \end{aligned} \tag{5.33}$$

We want $\delta(v^f) \leq \frac{1}{\sqrt{2}}$ this implies that $4\delta^2(v^f) \leq 2$. From (5.17) we have

$$\delta^2(v^f) \leq \frac{\frac{1}{4}\left(\frac{\left\|d_x^f\right\|^2+\left\|d_s^f\right\|^2}{1-\theta}\right)^2}{1-\frac{1}{2}\left(\frac{\left\|d_x^f\right\|^2+\left\|d_s^f\right\|^2}{1-\theta}\right)} \leq 2. \tag{5.34}$$

Let us set $u := \frac{\left\|d_x^f\right\|^2+\left\|d_s^f\right\|^2}{1-\theta}$ then we have

$$\begin{aligned} \tfrac{1}{4}u^2 &\leq 2(1-\tfrac{1}{2}u) \\ \tfrac{1}{4}u^2 &\leq 2-u \end{aligned}$$

$$\begin{aligned} \tfrac{1}{4}u^2 + u - 2 &\leq 0 \\ u^2 + 4u - 8 &\leq 0 \end{aligned}$$

Solving for u, we have

$$\begin{aligned} u_{1,2} &= \tfrac{-4\pm\sqrt{16+32}}{2} \\ &= -2 \pm \sqrt{12} \\ &\cong 1.46. \end{aligned}$$

Hence, we only choose the positive u-value. Thus, if $u := \frac{\|d_x^f\|^2+\|d_s^f\|^2}{1-\theta} \leq 1.46$, then inequality (5.34) holds. The condition was also that (5.34) holds if $\|d_x^f d_s^f\|_\infty \leq 1-\theta$ which implies

$$\begin{aligned} &\Rightarrow \tfrac{\|d_x^f d_s^f\|}{1-\theta} \leq 1 \\ &\Rightarrow \tfrac{1}{2}\tfrac{\|d_x^f\|^2+\|d_s^f\|^2}{1-\theta} \leq 1 \\ &\Rightarrow \tfrac{\|d_x^f\|^2+\|d_s^f\|^2}{1-\theta} \leq 2. \end{aligned}$$

Thus both conditions are satisfied, and we have

$$\|d_x^f\|^2 + \|d_s^f\|^2 \leq 1.46(1-\theta) \quad . \tag{5.35}$$

Combining equations (5.33) and (5.35) we get

$$(4\delta^2 + 2\theta n) + 18\theta^2 n^2 (2+q(\delta))^2 q^2(\delta) + 6\theta n\sqrt{4\delta^2+2\theta n}(2+q(\delta))q(\delta) \leq 1.46(1-\theta). \tag{5.36}$$

We know that old $\delta(x,s;\mu) \leq \tau \leq \frac{1}{\sqrt{2}}$. We also see that $q(\delta) = \delta + \sqrt{\delta^2+1}$ is increasing in δ and therefore the entire left side of equation (5.37) is increasing in δ. If we call $(2+q(\delta))q(\delta) = \bar{q}$, we have

$$4\delta^2 + 2\theta n + 18\bar{q}^2 n^2\theta^2 + 6\bar{q}n\theta\sqrt{4\delta^2+2\theta n} \leq 1.46(1-\theta). \tag{5.37}$$

Finally, we have to find τ and θ such that (5.37) is satisfied. The following table gives the answer. Explanations of tabular solutions: p_{soln} represents the solution of the left inequality of (5.37) and f_{soln} represents the solution of the right inequality of (5.37). The statement "f is less than p" implies that the θ parameter failed the inequality (5.37) for the specified τ value.

	τ	$\frac{1}{8}$		$\frac{1}{4}$		$\frac{1}{2}$	
θ		p_{soln}	f_{soln}	p_{soln}	f_{soln}	p_{soln}	f_{soln}
$\frac{1}{10n}$		1.0621274	1.314	f is less than p		f is less than p	
$\frac{1}{12n}$		0.8520225	1.3383333	1.3134135	1.3383333	f is less than p	
$\frac{1}{14n}$		0.7124501	1.3557143	1.1303083	1.3557143	f is less than p	

Tab.1, Choice of Parameters τ, θ

Thus, $\tau = \frac{1}{4}$ and $\theta = \frac{1}{12n}$ are the best chosen parameters that satisfy inequality (5.37) for any n, although practically for $n \geq 2$.

The above discussion can be summarized in the following theorem.

Theorem 5.1.6. *Let $\theta = \frac{1}{12n}$, $\tau = \frac{1}{4}$, and (x, s, μ) be the starting iteration with $\delta(x, s, \mu) = \delta$. Then after the feasibility step, we obtain (x^f, s^f) that are strictly feasible for P_{v^+} and $\delta(x^f, s^f, \mu^+) < \frac{1}{\sqrt{2}}$.*

5.2 Centering Step

After the feasibility step we have (x^f, s^f) feasible for LCP$_{\nu^+}$ such that $\delta(x^f, s^f; \mu^+) < \frac{1}{\sqrt{2}}$. Our next goal is to perform several centering steps to get sufficiently close to the μ^+-center of the LCP$_{\nu^+}$. Since $(x^f, s^f; \mu^+)$ is the starting iteration we denote them as $(x, s; \mu)$ respectively and we denote $\delta(x^f, s^f, ; \mu^+)$ as δ. To obtain our centering directon, we use the following system:

$$\begin{aligned} M\Delta x - \Delta s &= 0 \\ S\Delta x + X\Delta s &= \mu e - xs. \end{aligned} \tag{5.38}$$

Also, recall that

$$v = \sqrt{\frac{xs}{\mu}}\ ,\ \Delta x = \frac{x d_x}{v}\ ,\ \Delta s = \frac{s d_s}{v} \tag{5.39}$$

Then (5.38) becomes

$$\begin{aligned} \widetilde{M}dx - ds &= 0 \\ Sdx + Xds &= v^{-1} - v. \end{aligned} \tag{5.40}$$

Before we continue the analysis we give several helpful equations and inequalities.

Helpful Equations and Inequalities

From Chapter 4, we have that $\tilde{M}dx = ds$. Since $\tilde{M}$ is positive semi-definite, we can write $dx^T ds = dx^T \widetilde{M} dx \geq 0$, so $dx^T ds \geq 0$. Let $p_v = dx + ds$ and $q_v = dx - ds$ then

$$\begin{aligned} \|p_v\|^2 - \|q_v\|^2 &= p_v^T p_v - q_v^T q_v \\ &= (dx + ds)^T (dx + ds) - (dx - ds)^T (dx - ds) \\ &= 4dx^T ds. \end{aligned} \tag{5.41}$$

Since $dx^T ds \geq 0$, we conclude that $\|p_v\|^2 \geq \|q_v\|^2$. Furthermore, we can write

$$\begin{aligned} dx^T ds &\leq \tfrac{1}{4} \|p_v\|^2 \\ &= \tfrac{1}{4} \|dx + ds\|^2 \\ &= \delta^2. \end{aligned} \tag{5.42}$$

Therefore we have

$$0 \leq dx^T ds \leq \delta^2. \tag{5.43}$$

From the equation (5.41) we get

$$\begin{aligned} \|p_v\|^2 - \|q_v\|^2 &= 4dx^T ds \\ \|q_v\|^2 &= \|p_v\|^2 - 4dx^T ds \\ &\leq \|p_v\|^2, \quad \text{since } dx^T ds \geq 0 \\ &= \|dx + ds\|^2 \\ &= 4\delta^2. \end{aligned} \tag{5.44}$$

Now we consider $p_v^2 - q_v^2$:

$$\begin{aligned} p_v^2 - q_v^2 &= (dx - ds)^2 - (dx + ds)^2 \\ &= 4dxds \\ dxds &= \tfrac{1}{4}(p_v^2 - q_v^2) \\ |dxds| &= \tfrac{1}{4}\left|(p_v^2 - q_v^2)\right|. \end{aligned} \tag{5.45}$$

Case I : $p_v^2 - q_v^2 \geq 0$. Given that $p_v^2 - q_v^2 \leq p_v^2$ it follows that

$$\begin{aligned} |dxds| &= \tfrac{1}{4}\left|p_v^2 - q_v^2\right| \\ &= \tfrac{1}{4}(p_v^2 - q_v^2) \\ &\leq \tfrac{1}{4}p_v^2 \\ &= \tfrac{1}{4}\left|p_v\right|^2. \end{aligned} \tag{5.46}$$

Case II : $p_v^2 - q_v^2 \leq 0$. This implies $q_v^2 - p_v^2 \geq 0$, i.e. $|p_v^2 - q_v^2| \leq q_v^2$.

$$\begin{aligned} |dxds| &= \tfrac{1}{4}\left|p_v^2 - q_v^2\right| \\ &\leq \tfrac{1}{4}\left|q_v\right|^2. \end{aligned} \tag{5.47}$$

Thus,

$$\max_i |dx_i ds_i| \leq \tfrac{1}{4}\max\left\{|p_v|^2, |q_v|^2\right\},$$

which leads to

$$\max_i |dx_i ds_i| = \|dxds\|_\infty \leq \tfrac{1}{4}\max\left\{\|p_v\|^2, \|q_v\|^2\right\}$$

Therefore, from equation (5.44) we have,

$$\|dxds\|_\infty \leq \delta^2. \tag{5.48}$$

Next, we have

$$\begin{aligned} \|dxds\|^2 &= (dxds)^T(dxds) \\ &= (dx_1ds_1)^2 + (dx_2ds_2)^2 + ... + (dx_nds_n)^2 \\ &\leq (dx_1ds_1 + dx_2ds_2 + ... + dx_nds_n)^2 \\ &= (dx^Tds)^2 \\ &\leq \delta^4. \end{aligned} \tag{5.49}$$

Hence, $\|dxds\| \leq \delta^2$.

Now, we can easily obtain similar inequalities for Δx and Δs:

$$\begin{aligned} \Delta x^T \Delta s &= (xv^{-1}dx)^T(sv^{-1}ds) \\ &= (dx\sqrt{\tfrac{x}{s}}\sqrt{\mu})^T(ds\sqrt{\tfrac{s}{x}}\sqrt{\mu}) \\ &= \mu(dx\sqrt{\tfrac{x}{s}})^T(ds\sqrt{\tfrac{s}{x}}) \\ &= \mu dx^T ds \\ &\leq \mu\delta^2, \end{aligned} \tag{5.50}$$

$$\begin{aligned} \|\Delta x \Delta s\|_\infty &= max_i\,|\Delta x_i \Delta s_i| \\ &= max_i\,|\mu dx_i ds_i| \\ &= \mu max_i\,|dx_i ds_i| \\ &= \mu\,\|dxds\|_\infty \\ &\leq \mu\delta^2, \end{aligned} \tag{5.51}$$

$$\begin{aligned} \|\Delta x \Delta s\|^2 &= \textstyle\sum_{i=1}^n (\Delta x_i \Delta s_i)^2 \\ &= \textstyle\sum_{i=1}^n \mu^2 (dx_i ds_i)^2 \\ &= \mu^2 \textstyle\sum_{i=1}^n (dx_i ds_i)^2 \\ &= \mu^2\,\|dxds\|^2, \\ \|\Delta x \Delta s\| &= \mu\,\|dxds\| \\ &\leq \mu\delta^2. \end{aligned} \tag{5.52}$$

Continuing with the analysis we have that the new centering iterate is $x^+ = x + \Delta x$ and $s^+ = s + \Delta s$. Then,

$$\begin{aligned} x^+ s^+ &= (x+\Delta x)(s+\Delta s) \\ &= xs + x\Delta s + s\Delta x + \Delta x \Delta s \\ &= xs + (\mu e - xs) + \Delta x \Delta s \\ &= \mu e + \Delta x \Delta s. \end{aligned} \tag{5.53}$$

using equation (5.39) we have

$$\begin{aligned} x^+ s^+ &= \mu e + \tfrac{x d_x}{v} \tfrac{s d_s}{v} \\ &= \mu e + \mu dxds \\ &= \mu (e + dxds). \end{aligned} \tag{5.54}$$

Lemma 5.2.1. $(x^+)^T s^+ \leq \mu(n + \delta^2)$

Proof. We have

$$\begin{aligned} (x^+)^T s^+ &= e^T(x^+ s^+) \\ &= e^T(x+\Delta x)(s+\Delta s) \\ &= e^T(\mu e + \Delta x \Delta s) \\ &= \mu e^T e + e^T \Delta x \Delta s \\ &= \mu n + \Delta x^T \Delta s \\ &\leq \mu n + \mu \delta^2 \\ &= \mu(n + \delta^2). \end{aligned} \tag{5.55}$$

$\square$

The immediate consequence is the following corollary.

Corollary 5.2.2.

$$\|v\|^2 \leq n + \delta^2.$$

Proof. We have

$$\begin{aligned}
\|v\|^2 &= v^T v \\
&= \left(\sqrt{\tfrac{xs}{\mu}}\right)^T \left(\sqrt{\tfrac{xs}{\mu}}\right) \\
&= \tfrac{1}{\mu}(x_1 s_1 + x_2 s_2 + ... + x_n s_n) \\
&\leq \tfrac{1}{\mu}(\mu(n + \delta^2)) \\
&= n + \delta^2.
\end{aligned} \tag{5.56}$$

□

We know the initial iterate of the centering step is feasible because the requirement of the feasibility step is to get strictly feasible iterates for the full Newton step; therefore, $e + d_x d_s \geq 0$ or strictly feasible when $e + d_x d_s > 0$. Also from the feasibility step we know $\delta < \frac{1}{\sqrt{2}} < 1$.

Lemma 5.2.3. *If $\delta < 1$, then x^+ and s^+ are positive, i.e., they are strictly feasible and $\delta(x^+, s^+, \mu) \leq \frac{\delta^2}{2\sqrt{1-\delta^2}}$.*

Proof. Let $\delta^+ = \delta(x^+, s^+, \mu)$ and $v^+ = \sqrt{\frac{x^+ s^+}{\mu}}$. Since $\delta(v) = \frac{1}{2}\left\|v^{-1} - v\right\|$, we have

$$\begin{aligned}
\delta^+ &= \tfrac{1}{2}\left\|(v^+)^{-1} - v^+\right\| \\
&= \tfrac{1}{2}\left\|(v^+)^{-1}(e - (v^+)^2)\right\|.
\end{aligned} \tag{5.57}$$

From (5.53) we have $x^+ s^+ = \mu(e + dxds)$, and v^+ becomes $v^+ = \sqrt{e + dxds}$. By substituting this value into the equation (5.57) we get

$$\begin{aligned}
2\delta^+ &= \left\|(\sqrt{e+dxds})^{-1} - (e - e - dxds)\right\| \\
&= \left\|\tfrac{dxds}{\sqrt{e+dxds}}\right\| \\
&\leq \tfrac{\|dxds\|}{\left\|\sqrt{e+dxds}\right\|} \qquad (5.58) \\
&\leq \tfrac{\|dxds\|}{\sqrt{1-\|dxds\|_\infty}} \\
&\leq \tfrac{\delta^2}{\sqrt{1-\delta^2}}.
\end{aligned}$$

Thus,

$$\delta^+ \leq \frac{\delta^2}{2\sqrt{1-\delta^2}}$$

□

Corollary 5.2.4. *If* $\delta(x, s, \mu) \leq \frac{1}{\sqrt{2}}$ *then* $\delta(x^+, s^+, \mu) \leq \delta^2(x, s, \mu)$.

Proof. From Lemma 5.2.3 we have

$$\begin{aligned}
\delta(x^+, s^+, \mu) &\leq \tfrac{\delta^2(x,s,\mu)}{2\sqrt{1-\delta^2(x,s,\mu)}} \\
&\leq \tfrac{\delta^2(x,s,\mu)}{2\sqrt{1-\frac{1}{2}}} \qquad (5.59) \\
&= \tfrac{\delta^2(x,s,\mu)}{\sqrt{2}} \\
&\leq \delta^2(x, s, \mu),
\end{aligned}$$

which proves the corollary. □

Corollary 5.2.4 actually indicates that we have quadratic convergence if the iterates are sufficiently close to the μ center.

To determine the necessary number of centering steps, we use the fact that $\delta(x^+, s^+; \mu) \leq \delta^2$, and we continue with centering steps until $\delta(x^+, s^+; \mu) \leq \tau = \frac{1}{4}$.

How many do we need? Let k denote the number of centering steps. Then we have

$$
\begin{aligned}
\delta^{2k} &\leq 2^{-2} \\
\log \delta^{2k} &\leq \log 2^{-2} \\
2k \log \delta &\leq -2 \log 2 \\
2k \log \tfrac{1}{\sqrt{2}} &\leq -2 \log 2 \\
-k \log 2 &\leq -2 \log 2 \\
k &\geq 2
\end{aligned}
\tag{5.60}
$$

Therefore, we require only two centering steps per each feasibility step. Thus, all the iterates of the algorithm are guaranteed to be in the same neighborhood ($\tau = \frac{1}{4}$) of the central path. This leads to the following estimate on the number of iterations to obtain ϵ-approximate solution of the LCP.

Theorem 5.2.5. *If $\theta = \frac{1}{12n}$, $\mu_0 = \zeta_P \zeta_D$, and $\tau = \frac{1}{4}$ then the Infeasible Full Newton-step IPM requires at most $12nlog\frac{33(x^0)^T(s^0)}{32\epsilon}$ iterations to obtain ϵ-approximate solution of LCP(M,q) or equivalently $O(n \log \frac{n}{\epsilon})$ iterations.*

Proof. At the start of the algorithm, the duality gap has a certain value and in each iteration the duality gap is reduced by the factor $1 - \theta$. The duality gap can be transformed as follows

$$
\begin{aligned}
x_k^T s_k &\leq \mu_k(n + \delta^2) \\
&\leq \mu_k(n + \tfrac{1}{16}) \\
&= (1-\theta)^k \mu_0 (n + \tfrac{1}{16}) \\
&= (1-\theta)^k \zeta_P \zeta_D (n + \tfrac{1}{16}) \\
&\leq \zeta_P \zeta_D (n + \tfrac{1}{32} n) \\
&\leq (1-\theta)^k (x^0)^T (s^0) \tfrac{33}{32} \leq \epsilon
\end{aligned}
$$

This is satisfied if

$$\begin{aligned}
\log[(1-\theta)^k (x^0)^T (s^0) \tfrac{33}{32}] &\leq \log \epsilon \\
\log(1-\theta)^k + \log(x^0)^T (s^0) + \log \tfrac{33}{32} &\leq \log \epsilon \\
\log(1-\theta)^k &\leq \log \epsilon - \log(x^0)^T (s^0) - \log \tfrac{33}{32} \\
k \log(1-\theta) &\leq \log \epsilon - \log(x^0)^T (s^0) - \log \tfrac{33}{32} \\
-k \log(1-\theta) &\geq -\log \epsilon + \log(x^0)^T (s^0) + \log \tfrac{33}{32}
\end{aligned}$$

since $-log(1-\theta) \geq \theta$, we have

$$\begin{aligned}
k\theta &\geq \log \tfrac{(x^0)^T (s^0) \frac{33}{32}}{\epsilon} \\
k &\geq \tfrac{1}{\theta} \log \tfrac{(x^0)^T (s^0) \frac{33}{32}}{\epsilon} \\
k &\geq \tfrac{1}{\frac{1}{12n}} \log \tfrac{(x^0)^T (s^0) \frac{33}{32}}{\epsilon} \\
k &\geq 12n \log \tfrac{33 (x^0)^T (s^0)}{32\epsilon}
\end{aligned}$$

concludes the number of iterations needed for the feasibility step. Since we need two centering steps per each feasibility step, the total number of iterations needed is $k \geq 2 \times 12n \log \frac{33(x^0)^T(s^0)}{32\epsilon}$ which equals $24n \log \frac{33(x^0)^T(s^0)}{32\epsilon}$. It is easy to see that the number of iterations are $O(n \log \frac{n}{\epsilon})$.

CHAPTER 6

NUMERICAL RESULTS

In this chapter, the Infeasible Full Newton-step interior-Point Algorithm for LCP, as given in Table 4.1, is implemented in MATLAB. We performed numerical tests of our implementation of the algorithm for a set of problems of various dimensions. Some problems were generated "by hand" and others were randomly generated. The summary of results is given in tables below. Note that for all tables below, excluding table 2, $\zeta_P = 1$ and $\zeta_D = 1$.

6.1 Generating Sample Problems

Generating Matrix M

Before we go into the numerical data, we briefly describe how the test problems were generated. The first group of problems were manually generated. The PSD matrices of the problems were generated by using "rand" function as described below.

$$\begin{aligned} A &= rand(k, n), \text{ where } 1 \leq k \leq n \\ M &= A^T A. \end{aligned}$$

Starting points and initial conditions

To initiate the progress we first choose x^0 and s^0 as vectors of ones. We also examine the more general case $x^0 = \zeta_P e, s^0 = \zeta_D e$ for some parameter $\zeta_p, \zeta_D > 0$. For our first testing, we set $\zeta_P = 1$ and $\zeta_D = 1$, and we take the set of parameters $\tau = \frac{1}{4}$ and $\theta = \frac{1}{12n}$ as required by the algorithm in order to guarantee convergence. After the test for $\zeta' s$ of the same value, we test different ζ values to show the efficiency of the algorithm converging from different (feasible/infeasible) starting points.

Parameters

As will be shown, several sets of parameters were tested. We take action on a τ-neighborhood ($\tau = \frac{1}{4}$) and more aggressive reduction of μ-parameter at each iteration, by taking the barrier parameter to be a fixed value independent of the size of the problem ($\theta = \frac{1}{\sqrt{12n}}$). In this case we can not guarantee convergence, however, in most instances the algorithm still converges.

Finally, we try a wider τ-neighborhood ($\tau = \frac{1}{3}$) and more aggressive reduction of μ-parameter at each iteration, by taking the barrier parameter to be a fixed value independent of the size of the problem ($\theta = \frac{1}{\sqrt{6n}}$). This case again, does not guarantee convergence, however, in most instances the algorithm still converges.

6.2 Summary of Numerical Results

We generated 9 test problems. Two of them were generated manually (denoted as EH) with dimensions up to $n = 5$ and seven were randomly generated (denoted ER) with dimensions up to $n = 200$. This set of test problems were solved with the following set of parameters.

1. $\tau = \frac{1}{4}$ and $\theta = \frac{1}{12n}$,
2. $\tau = \frac{1}{4}$ and $\theta = \frac{1}{\sqrt{12n}}$,
3. $\tau = \frac{1}{3}$ and $\theta = \frac{1}{\sqrt{6n}}$

The number of iterations as well as CPU time (CPU time in seconds) for each case are listed in the tables below.

$Problem$	$Size$	$CPUtime$	$Iterations$
$EH1$	3×3	$1.9912e^{-2}$	374
$ER2$	3×3	$1.6756e^{-2}$	366
$EH3$	5×5	$2.102e^{-2}$	644
$ER4$	5×5	$2.4774e^{-2}$	709
$ER5$	5×5	$2.5357e^{-2}$	714
$ER6$	10×10	$7.082e^{-2}$	1609
$ER7$	50×50	1.919146	10559
$ER8$	100×100	12.953707	23197
$ER9$	200×200	139.182232	50548

Tab.1, $\theta = \frac{1}{12n}, \ \tau = \frac{1}{4}$

To further show the efficiency of the algorithm, we also tested example EH1 with different μ starting point values, where again $\mu = \zeta_P \zeta_D$. Hence we have the following:

$Problem$	ζ_P	ζ_D	$Size$	$CPUtime$	$Iterations$
$EH1$	2	3	3×3	$2.4180e^{-2}$	430
$EH1$	7	15	3×3	$2.7318e^{-2}$	532
$EH1$	100	48	3×3	$3.2984e^{-2}$	667

Tab.2, $\theta = \frac{1}{12n}, \ \tau = \frac{1}{4}$

Table 2 shows us that no matter where our starting point is, rather it be feasible or infeasible, the IIPM Algorithm will converge on these problems. The number of iterations slightly increase; however, the algorithm converged to the same solution.

Note: We would also like to point out that the solution of EH1 obtained using our IPM matches the solution obtained from using the classical Lemke's algorithm. This is a strong indicator of the correctness of our implementation of IIPM.

In the Table 3 below we fixed the barrier update parameter θ for all the examples and we did not change the threshold parameter. So $\tau = \frac{1}{4}$ and $\theta = \frac{1}{\sqrt{12n}}$. Though this update of $\theta = \frac{1}{\sqrt{12n}}$ does not guarantee convergence, if it converges, the convergence is much faster.

$Problem$	$Size$	$CPUtime$	$Iterations$
$EH1$	3×3	$4.471e^{-3}$	58
$ER2$	5×5	$3.102e^{-3}$	83
$ER3$	5×5	$6.087e^{-3}$	86
$ER4$	10×10	$8.689e^{-3}$	144
$ER5$	50×50	$8.9837e^{-2}$	423
$ER6$	100×100	$3.73567e^{-1}$	661
$ER7$	200×200	2.949483	1023

Tab.3, $\theta = \frac{1}{\sqrt{12n}}, \tau = \frac{1}{4}$

Although the convergence is not guaranteed, we see that the algorithm still converges for all test problems, and that there is a significant reduction of CPU time as well.

In the Table 4 below we increase both the threshold parameter and the barrier update parameter. So $\tau = \frac{1}{3}$ and $\theta = \frac{1}{\sqrt{6n}}$.

$Problem$	$Size$	$CPUtime$	$Iterations$
$EH1$	3×3	$5.816e^{-3}$	57
$ER2$	3×3	$4.642e^{-3}$	39
$ER3$	5×5	$5.082e^{-3}$	59
$ER4$	10×10	$7.377e^{-3}$	98
$ER5$	50×50	$9.1837e^{-2}$	297
$ER6$	100×100	$2.65103e^{-1}$	464
$ER7$	200×200	2.089771	720

Tab.4, $\theta = \frac{1}{\sqrt{6n}}, \tau = \frac{1}{3}$

The increase in the paramaters θ and τ leads to a further reduction of iterations and CPU time. Even this preliminary implementation shows that the method is computationally competitive with IPM methods that require calculation of a step size and very often feasbility condition for the starting point. Both of these conditions are not required here.

CHAPTER 7

CONCLUSION

In this thesis, we consider the Monotone Linear Complementarity Problem (LCP) defined by (2.1) with positive semidefinite matrix. Although LCP is not an optimization problem, it is closely related to many important optimization problems and it has many important applications.

The LCP problem can be solved using classical simplex-type (pivoting) Lemke's algorithm that is described in Chapter 3. However in the last two decades a new class of Newton-type IPM have been developed and successfully applied to solve LCP.

We propose a new IPM to solve the Monotone LCP. The algorithm is given in Table 4.1 in Chapter 4. There are two main features of the IPM. First is that there is no calculation of a step-size, i.e., we use full Newton step at each iteration. The second feature is that we can start from any point. This point may or may not be feasible, and that is why we call the algorithm Infeasible Full-Newton-Step IPM (IIMP). We show that the convergence of the algorithm is guaranteed by appropriate choice of parameters θ (barrier parameter) and τ (threshold parameter). We prove that if $\theta = \frac{1}{12n}$ and $\tau = \frac{1}{4}$ then the iteration bound is $O(n \log \frac{n}{\epsilon})$ which matches the best known iteration bound for these types of methods. This convergence analysis is the emphasis of the thesis and it is provided in Chapter 5.

If θ depends on n such as in our algorithm, $\theta = O(\frac{1}{n})$, then the algorithm call a short-step algorithm. If θ is independent of n such as $\theta = O(1)$, then the algorithm is called a long-step algorithm. In our method, in order to prove convergence result, parameter θ depends on n, therefore the method is a short-step algorithm.

Furthermore, in Chapter 6 we also provided an initial implementation of the method and tested it on a small set of test problems of various dimensions and various starting points. Several sets of parameters were tested in the implementation of the algorithm. First we set the parameter to be $\tau = \frac{1}{4}$ and $\theta = \frac{1}{12n}$ as required by the algorithm in order to guarantee convergence. Next, we maintained all parameters and changed the starting position by changing the starting points. Lastly, we tried a wider τ-neighborhood, $\theta = \frac{1}{\sqrt{6n}}$, $\tau = \frac{1}{3}$ and this yielded more aggressive reductions of μ-parameter at each iteration and quicker convergence;(the number of itererations reduced significantly) however, in general, convergence in this case may not be achieved.

The results we obtained show that the method converges for all test problems even in the case when choice of parameters does not theoretically guarantee convergence. Now, the initial implementation although not sophisticated still show promising results. Even more importantly, the fact that convergence was reached no matter the starting point shows the robustness of the algorithm. Though the number of iterations increased somewhat with respect to greater infeasible starting points, the increase in CPU time was minimal. Overall, the proposed algorithm is both theoretically and practically promising.

APPENDIX A

MATLAB Codes

The following is the listing of the main program of the Infeasible Full Newton Step IPM for LCP. The input data is generated either randomly or manually. The results are yield by using the subroutine IPMtre.m which implements the Algorithm to solve the problem.

A.1 Main Program : *mainIPM.m*

```
%program Thesis
clc
clear;
tic
epsilon=10^-4;

%The following Load commands inputs the data to be
%calculated by hand.

% load M.txt
% load q.txt

% For the Matrices generated by using random generater (for the second set of examples)
%choose n>=2

A = rand(k,n);
%k is in the interval [1,n]
%We create a positive semidefinite matrix M from matrix A
M = A'A;
n=length(M);
q=rand(n,1);
eig(M);

% theta=1/(12*n);
% theta=1/sqrt(12*n);
theta=1/sqrt(6*n);
% tau=1/4
tau=1/3;

[x s] = IPMtre(M,n,epsilon,theta,tau,q);

toc
```

A.2 IIPM Algorithm : *IPMtre.m*

```
function [x,s]= IPMtre(M,n,epsilon,theta,tau,q)

zetap=1;
zetad=1;
x=(zetap)*ones(n,1);
s=(zetad)*ones(n,1);
mu=(zetap)*(zetad);
nu=1;

%outer loop
r=(-M*x-q+s);
count1=0;
count2=0;
ru=nu*r;

while max(x'*s,norm(ru))>=epsilon
    count1=count1+1;
    v=sqrt(x.*s./mu);

    X = diag(x);
    S = diag(s);

    Dx = (S+ X*M)\( X*theta*nu*r + (1-theta)*mu*ones(n,1) -X*S*ones(n,1));
    Ds = M*Dx - theta*nu*r;
    x=x+Dx;
    s=s+Ds;
    mu=(1-theta)*mu;
    nu=(1-theta)*nu;
    v=sqrt(x.*s./mu);

    delta=norm(v.^(-1) - v)/2;
    ru=nu*r;

    %inner loop
    count2=0;
    while delta >= tau
        count2=count2+1;

        Dx = (S+ X*M)\(mu*ones(n,1)-X*S*ones(n,1));
        Ds = M*Dx ;

        x=x+Dx;
        s=s+Ds ;
        v=sqrt(x.*s./mu);
        delta=0.5*norm(v-v.^(-1));

    end
```

```
end
 count1
 count2
end
```

APPENDIX B

MATLAB Output

Below we provide the entire output for Example 3.3.

B.1 *Output of EH1*

```
  % Mathlab output of the example describe in both Lemke's method and the Infeasible IPM
count1 =

   374


count2 =

     0


x =

    0.0000
    2.0000
    1.0000


s =

    1.0000
    0.0000
    0.0000

Elapsed time is 0.019912 seconds.
```

BIBLIOGRAPHY

[1] Bai Y., Roos C., *A New Class of Polynomial Interior-Point Algorithms for Linear Complementarity Problems*, Pacific Journal of Optimization, Vol. 4, No. 1, (2008).

[2] Bai Y., Lesaja G., Roos C., Wang G., Ghami M. El, *A Class of Large and Small Update Primal-Dual Interior-Point Algorithms for Linear Optimization*, Journal of Optimization Theory and Applications, Vol. 138, No. 3, pp. 341-359, (2008).

[3] Cottle R., Pang J., Stone R., *The Linear Complementarity Problem*, Academic Press, Inc., Boston, (1992).

[4] Cottle and Dantzig (1968), *Complementary pivot theory of mathematical programming*, Linear Algebra and Its Applications 1.

[5] Facchinei F., Pang J.S., *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer, New York, (2003)

[6] Karmarkar N., *A New Polynomial Time Algorithm for Linear Programming*, Combinatorica, Vol 4, No. 4, (1984).

[7] Khachiyan L.G., *A Polynomial Algorithm in Linear Programming*, Soviet Mathematics Doklady, 20, pp. 373-395 (1984).

[8] Klee V. and Minty G.J., *How Good is the Simplex Algorithm?* Inequalities, III, pp. 159-175, Academic Press, New York, NY, (1972).

[9] Kojima M., Megiddo N., Noma T., Yoshise A., *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems*, Springer-Verlag, Berlin, Germany (1991).

[10] Lemke C.E., *Bimatrix Equilibrium Points and Mathematical Programming*, Management Science II, pp. 681-689, (1965).

[11] Lesaja G., *Introducing Interior-Point Methods for Introductory Operations Research Courses and/or Linear Programming Courses*, Open Operational Research Journal, Vol. 3, pp. 1-12, (2009).

[12] G. Lesaja and C. Roos. Unified analysis of kernel-based interior-point methods for $P_*(\kappa)$-LCP. *SIAM Journal on Optimization,* 20(6): 3014-3039, 2010.

[13] Lesaja G., Mansouri H., Roos C.,Zangiabadi M., *Full-Newton-Step Interior-point Methods for Linear Optimization Based on Locally Self-Concordant Barrier Functions*, Mathematical Programming, accepted, (2010).

[14] Mansouri H., *Full-Newton-Step Interior-point Methods for conic Optimization*, Ph.D. Thesis, TU Delft, Netherland, (2008).

[15] Nesterov Y., Nemirovski A., *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM Studies in Applied Mathematics, Philadelphia, PA, (1994).

[16] Megiddo N., *Pathways to the Optimal Set in Linear Programming*, Progress in Mathematical Programming: Interior Point and Related Methods, pp. 131-158, Springer, New York, (1989).

[17] Peng J.S., Roos C., Terlak T., *Self-Regularity: A New Paradigm for Primal-Dual Interior-Point Algorithms*, Princeton University Press, (2002)

[18] Roos C., *A Full-Newton Step $O(n)$ Infeasible Interior-Point Algorithm for Linear Optimization*, *SIAM Journal on Optimization,* 16(6): 1110-1136, 2006.

[19] Roos C., Terlaky T., Vial J.P., *Theory and Algorithms for Linear Optimization*, John Wiley and Sons, Chichester, UK (1997).

[20] Shor N.Z., *Cut-off Method with Space Extension in Convex Programming Problems*, Cybernetics 13, pp. 94-96, (1977).

[21] Sonnevend G., *An "analytic center" for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming*, System Modeling and Optimization. Proceedings of the 12th IFIP-Conference, Budapest, Hungary, September 1985. Lecture Notes in Control and Information Sciences, vol. 84, pp. 866-876. Springer, Berlin (1986).

[22] Wright S., *Primal-Dual Interior-Point Methods*, SIAM Publishing, Philadelphia, PA (1997).

[23] Wright S., Ferris M., Mangasarian O., *Linear Programming with MATLAB*, SIAM, Series on Optimization, (2007).

[24] Nemirovski A., Yudin D.B., *Informational Complexity and Effective Methods of Solution for Convex Extremal Problems*, Ekonomika i Matematicheshie Metody 12 (in Russian), pp. 357-369, (1976).