Some Generalizations of The Linear Complementarity Problem

by

Peng Yi

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Engineering - Industrial Engineering and Operations Research

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ilan Adler, Chair
Professor Shmuel S. Oren
Professor Bernd Sturmfels

Spring 2016

Some Generalizations of The Linear Complementarity Problem

Abstract

Some Generalizations of The Linear Complementarity Problem

by

Peng Yi

Doctor of Philosophy in Engineering - Industrial Engineering and Operations Research

University of California, Berkeley

Professor Ilan Adler, Chair

In this thesis, we study two generalizations of the classical linear complementarity problem (LCP) - the weighted extended linear complementarity problem (wXLCP) and the complementarity problem (CP) over a general closed cone. Our goal is twofold: extend some fundamental results of the LCP to a more general setting and identify a class of nonmonotone problems which could be solved numerically. The thesis is organized as follows:

- In Chapter 1, we formulate problems relevant to our study and introduce background material that will be needed in the rest of the thesis.
- In Chapter 2, we formulate the weighted extended linear complementarity problem (XLCP), which naturally generalizes the LCP, the horizontal linear complementarity problem (HLCP) and the extended linear complementarity problem (XLCP). Motivated by important roles played by matrix theoretic properties in the LCP theory, we study the monotonicity, sufficiency, P-property and R_0-property in the setting of the XLCP. Together with two optimization reformulations of the problem, we establish several fundamental results. Specifically, we show that the characterizing conditions of the row and column sufficiency properties in [17] can be similarly described in the context of the wXLCP. Under the monotonicity property, the wXLCP is equivalent to a convex optimization problem and it is solvable whenever it is strictly feasible. Also, we show that the row sufficiency property ensures that every stationary point of some unconstrained optimization problem is a solution of the wXLCP.
- In Chapter 3, we propose the general notion of uniform nonsingularity property for transformations over Euclidean spaces. We show that this property is closely related to a number of existing properties in the literature. In particular, the variants of the uniform non-singularity property recover P-property of a matrix, the strong monotonicity of a nonlinear transformation, and a weaker version of the Cartesian P-property [3] and the P-type property in [8]. Also, we show that a form of this property implies the Lipschitzian GUS-property.

- In Chapter 4, we present the applications of the uniform non-singularity property to the solution property of complementarity problems over general closed convex cones. With the help of the barrier based smoothing approximation of the normal-map formulation of complementarity problems, we develop a homotopy path whose accumulation point is a solution of the problem. Moreover, we show that the path is convergent and every solution of the complementarity problem comes from the limit of the path, whence establishing the uniqueness of solution.

To my family

Contents

Acknowledgments

This thesis is a celebration of my cordial relationship with my advisor Professor Ilan Adler. I have benefited intellectually from his life experiences, piercing insights and tastes of research. His perpetual encouragement has illuminated my graduate life, and his kindness and continuous support allowed me to strike a good balance between research and family. This thesis would never have been possible without his kind support.

I would like to thank my thesis committee members Professor Shmuel S. Oren and Professor Bernd Sturmfels. Their time and efforts to improve my work on short notice are greatly appreciated.

I am extremely indebted to Professor Chek Beng Chua from Nanyang Technological University, Singapore. He introduced to me the field of complementarity problem and taught me how to conduct research before I came to Berkeley to pursue my graduate degree. At my crossroad, he wrote "I want you to know that it is difficult times like this that makes a man. I'm confident that with experiences like this, once you and your family get through this tough period, you'll be much more capable in handling whatever else life throws at you." He has also contributed to Chapter 3 and 4 of this thesis.

I thank all other IEOR graduate students in my cohort and friends at Berkeley for making my graduate life a much more joyful experience. Their presence reminded me that I am not alone and motivated me to move forward.

Finally, I wish to express my heartfelt gratitude and love to my wife, Lijun, and my daughter, Olivia. Their unconditional support and love helped me overcome the most depressing moment of my graduate study. I am happy that we made it together as a family.

Chapter 1

Background

This chapter is expository in nature, and its purpose is to formulate problems relevant to our study, and to introduce background material that will be needed in the rest of the thesis.

1.1 Introduction

A *complementarity problem*, abbreviated CP, is a problem of finding a vector in a finite-dimensional real vector space that satisfies a certain system of inequalities and equations. It is closely related to finite-dimensional optimization and is a powerful machinery for the modelling of equilibria of many kinds. As a result of its abroad range of associations and diverse applications, it has received wide attention since the mid-1960s. In a span of five decades, a great deal of research effort was devoted to the development of the subject, which now has become a well-established and fruitful discipline in the field of mathematical programming. Major aspects of those developments include new mathematical theory, a rich body of effective solution algorithms, a multitude of interesting connections to numerous disciplines and a wide range of applications in engineering and economics. The literature of CPs has benefitted from contributions made by pure and applied mathematicians, operations researchers, computer scientists, economists and engineers of many kinds. A comprehensive treatment of the subject can be found in the classical monographs [43, 11, 41].

Mathematically, many fundamental equations in engineering and economics are often described by a complementary relation between two sets of variables. Consequently, it is not surprising that many physical and economic equilibrium problems can be profitably formulated as CPs. For example, the renowned Walrasian law of competitive equilibria of exchange economies [62] can be formulated as a nonlinear complementarity problem in the price and excess demand variables. Popularized by the Lemke-Howson algorithm [29], linear complementarity problems are instrumental for the discovery of superb construction tool for the computations of equilibria of bimatrix games. Another example is the Wardrop principle of user equilibrium traffic theory [63] which has a natural formulation as a nonlinear complementarity system. We refer the reader to the excellent survey paper [13] and references therein for a list of applications of the complementarity problems known to date.

To motivate our discussion, we begin with the classical *linear complementarity problem* (LCP). Given a matrix $M \in \mathbb{R}^{n\times n}$ and a vector $q \in \mathbb{R}^n$, the linear complementarity problem is to find a vector $x \in \mathbb{R}^n$ such that

$$\mathbf{LCP}(M,q): \quad x \geq 0, \quad Mx+q \geq 0 \quad \text{and} \quad \langle x, Mx+q\rangle = 0, \tag{1.1}$$

where $\langle x, y\rangle \triangleq \sum_{i=1}^n x_i y_i$ denotes the inner product between x and y. Despite of the simple form, LCPs cover many classes of problems, from well understood and easy problems to NP-hard problems, see for example, [43, 41]. Historically, one particularly important and well known context in which the LCP was found is the optimality conditions of linear and quadratic programs. Consider the following quadratic programming (QP):

$$\begin{aligned} \text{minimize} \quad & c^T x + \frac{1}{2} x^T Q x \\ \text{subject to} \quad & Ax \geq b, \\ & x \geq 0. \end{aligned} \tag{1.2}$$

where $Q \in \mathbb{R}^{n\times n}$ is symmetric, $c \in \mathbb{R}^n, A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$. Let x be a local optimal solution of (1.2), the Karush-Kuhn-Tucker (KKT) conditions imply that there exists $y \in \mathbb{R}^m$ such that

$$\begin{aligned} u = c + Qx - A^T y \geq 0, \quad x \geq 0, \quad x^T u = 0, \\ v = -b + Ax \geq 0, \quad y \geq 0, \quad y^T v = 0. \end{aligned} \tag{1.3}$$

If Q is monotone, i.e., $x^T Q x \geq 0$ for all $x \in \mathbb{R}^n$, then the objective function of (1.2) is convex and conditions (1.3) are therefore sufficient for x to be globally optimal. It is easy to verify that conditions in (1.3) define the $\mathbf{LCP}(M,q)$ where

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

It follows that the KKT conditions of any quadratic program can be formulated as a LCP. Several effective algorithms for solving quadratic programs are based on their LCP formulations.

Much of the theory of the LCP as well as many numerical algorithms for its solutions are based on the assumption that the underlying matrix M belongs to a particular class of matrices. A great deal of research effort was devoted to investigating the relevant matrix classes, to examining their interconnections and to exploring their relationship to the LCP. Typically, these investigations can be classified into two major directions: The first one was related to the intrinsic properties of the matrix itself (e.g., monotone or sufficient matrices); the second one was related to properties of the corresponding LCP solution set (e.g., matrices that ensure the existence or uniqueness of solutions). The ultimate goal was to find the connections between the two sets of characterizations.

Generalizations of LCPs

In this subsection, we introduce a few generalizations of LCPs.

The Horizontal Linear Complementarity Problem

Given $M, N \in \mathbb{R}^{n\times n}$ and $q \in \mathbb{R}^n$, the *horizontal linear complementarity problem* (HLCP) is the problem of finding $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ such that

$$\textbf{HLCP}(M, q): \quad x,\ y \geq 0, \quad Mx - Ny = q \quad \text{and} \quad \langle x, y\rangle = 0. \tag{1.4}$$

Clearly, (1.4) reduces to the system (1.1) if the $N = I$ is the identity matrix. The term HLCP was coined in [43] and has become important in the study of feasible and infeasible interior point algorithms for linear and convex quadratic programs. As pointed out by Zhang [67] and Guler [20], the HLCP formulation of convex programs is better suited for computational aspects than the LCP formulation. The HLCP arises in many contexts since a result of Evaes and Lemke [10] shows that any piecewise linear system can be formulated as a HLCP.

The Extended Linear Complementarity Problem

Given $M, N \in \mathbb{R}^{m\times n}$ and a polyhedral $\mathcal{P} \subseteq \mathbb{R}^m$, the *extended linear complementarity problem* (XLCP) is the problem of finding $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ such that

$$\textbf{XLCP}(M, N, \mathcal{P}): \quad x,\ y \geq 0, \quad Mx - Ny \in \mathcal{P} \quad \text{and} \quad \langle x, y\rangle = 0. \tag{1.5}$$

This problem was introduced by Mangasarian and Pang in [34], where some basic properties were established with the aid of a bilinear program. Gowda [17] showed that the XLCP is equivalent to the generalized LCP of Ye [65]: given matrices $A, B \in \mathbb{R}^{m\times n}, C \in \mathbb{R}^{m\times k}$ and a vector $q \in \mathbb{R}^m$, find vectors x, y and z satisfying

$$\begin{aligned} x, y, z &\geq 0, \\ Ax + By + Cz &= q, \\ \langle x, y\rangle &= 0. \end{aligned}$$

It is evident that both the LCP and the HLCP are special cases of the XLCP. In addition, the vertical linear complementarity problem (VLCP) and the mixed linear complementarity problem (MLCP) can all be formulated as XLCPs [17].

The Weighted Horizontal Linear Complementarity Problem

Given $M, N \in \mathbb{R}^{n\times n}$ and $q \in \mathbb{R}^n$, the *weighted horizontal linear complementarity problem* (wHLCP) is the problem of finding $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ such that

$$\textbf{wHLCP}(M, N, \mathcal{P}): \quad x,\ y \geq 0, \quad Mx - Ny \in \mathcal{P} \quad \text{and} \quad x \circ y = w. \tag{1.6}$$

where the notation $\circ$ denotes the Hadmard product of two vectors, i.e., $x \circ y$ is the vector whose ith component is equal to $x_i y_i$ for all $i \in [n] \triangleq \{1, ..., n\}$. The wHLCP, introduced in [45], was motivated by modelling the Fisher market equilibrium problem. In the Fisher's model the market is composed of producers and consumers. Consumer i has a budget $w_i > 0$ to spend on buying goods from the producers in such a way that an individual utility function is maximized. The price equilibrium is an assignment of prices to goods so that when every consumer buys a maximal bundle of goods then the market clears, meaning that all the money is spent and all the goods are sold. Under linear utility assumptions, Eisenberg and Gale proved that the market-clearing prices are given by the optimal Lagrange multipliers of some convex optimization problem. Potra [45] showed that the KKT conditions for the Eisenberg and Gale formulation, which can be modelled as a nonlinear complementarity problem, can also be formulated as an instance of the wHLCP; and the latter can be solved more efficiently by interior-point algorithms.

In Chapter 2, we study the *weighted extended linear complementarity problem* (wXLCP) which includes the LCP, the HLCP, the XLCP and the wHLCP as special cases. We extend a number of fundamental results in the theory of LCP to the setting of the wXLCP. The motivation of this study is that we expect more economic and engineering equilibrium problems could be profitably formulated as an instance of the wXLCP.

Complementarity Problems over General Closed Convex Cones

We now consider complementarity problems over a general closed convex cone. Let $\mathbb{E}$ be a finite-dimensional Euclidean space equipped with inner product $(x, y) \in \mathbb{E} \times \mathbb{E} \mapsto \langle x, y\rangle \in \mathbb{R}$, and let $\mathcal{K} \subseteq \mathbb{E}$ is a closed convex cone with $\mathcal{K}^*$ denoting its dual cone, i.e.,

$$\mathcal{K}^* \triangleq \{y \in \mathbb{E} : \langle x, y\rangle \geq 0 \ \forall x \in \mathcal{K}\}.$$

Given a continuous transformation $F : \mathbb{E} \to \mathbb{E}$, the complementarity problem consists in finding a vector $x \in \mathbb{E}$ such that

$$\mathbf{CP}(F, \mathcal{K}) : \quad x \geq_{\mathcal{K}} 0, \quad F(x) \geq_{\mathcal{K}^*} 0 \quad \text{and} \quad \langle x, F(x)\rangle = 0. \tag{1.7}$$

We use the Löwner partial order $x \geq_{\mathcal{K}} y$ (respectively, $x >_{\mathcal{K}} y$) to mean $x - y \in \mathcal{K}$ (respectively, $x - y \in \operatorname{int}(\mathcal{K})$, the topological interior of $\mathcal{K}$). Complementarity problems come in various types with different choices of triplets $(\mathbb{E}, \mathcal{K}, F)$. In particular, if $\mathcal{K}$ is a *symmetric cone*, that is, $\mathcal{K}$ is both *homogeneous* and *self-dual* (see more details in Section 1.3), then (1.7) can be simplified as the *symmetric cone complementarity problem* (SCCP):

$$\mathbf{SCCP}(F, \mathcal{K}) : \quad x \geq_{\mathcal{K}} 0, \quad F(x) \geq_{\mathcal{K}} 0 \quad \text{and} \quad \langle x, F(x)\rangle = 0. \tag{1.8}$$

The SCCP subsumes the following familiar examples.

- When $\mathbb{E} = \mathbb{R}^n$, the space of n-dimensional vectors with the standard inner product $\langle x, y\rangle = x_1 y_1 + \cdots + x_n y_n$ and $\mathcal{K} = \mathbb{R}^n_+ \triangleq \{x \in \mathbb{R}^n : x \geq 0\}$ is the cone of vectors with nonnegative entries, we get the classical *nonlinear complementarity problem* (NCP).

- When $\mathbb{E} = \mathbb{R} \times \mathbb{R}^{n-1}$ with the inner product $\langle x, y \rangle = x_1 y_1 + \cdots + x_n y_n$ and the cone $\mathcal{K} = \mathcal{L}^n_+ = \{(x_1, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{n-1} : x_1 \geq \|\bar{x}\|\}$ is the second order cone, we obtain the *second-order cone complementarity problem* (SOCCP).
- When $\mathbb{E} = \mathbb{S}^n$, the space of real symmetric $n \times n$ matrices equipped with the inner product $\langle x, y \rangle = \mathrm{tr}(xy) \triangleq \sum_{i=1}^n (xy)_{ii}$ and $\mathcal{K} = \mathbb{S}^n_+$ is the cone of positive semidefinite matrices, we obtain the *semidefinite complementarity problem* (SDCP).

The CPs naturally arise as the KKT conditions of conic optimization problems. In particular, the NCP is closely related to the following nonlinear program

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & g(x) \leq 0, \quad x \geq 0. \end{array}$$

where $\theta : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable real-valued function and $g : \mathbb{R}^n \to \mathbb{R}^m$ is a continuously differentiable vector-valued function. In addition, the NCP provides a unified formulation for many equilibrium problems such as the traffic equilibrium problem, the spatial price equilibrium problem, and the Nash-Cournot equilibrium problem, see, e.g., [13].

Current research on numerical algorithms for solving $\mathbf{SCCP}(F, \mathcal{K})$ mainly focus on monotone problems, that is, F is a monotone transformation. It is of theoretic interest to study some non-monotone property for transformations over Euclidean spaces and analyze its applications to complementarity problems. In Chapter 3, we introduce the notion of *uniform non-singularity* (UNS) property and connect it to existing properties in the literature. In Chapter 4, we show that the UNS property allows us to apply the homotopy technique to find a solution to $\mathbf{CP}(F, \mathcal{K})$. Moreover, we show that it ensures the uniqueness of the solution.

1.2 Equivalent Formulations

Complementarity problems admit a number of equivalent formulations, many of which not only provide insights into the CPs, but also form the basis for the development of different methods for their solutions. In what follows, we briefly introduce the quadratic programming formulation for the LCP and the *natural map* and *normal map* formulations for a general CP.

Quadratic Programming Formulation

As mentioned earlier, the KKT conditions of any quadratic programming can be formulated as a LCP. On the other hand, the cornerstone for existence results of the $\mathbf{LCP}(M, q)$ is the associated quadratic program reformulation:

$$\begin{array}{lll} \mathbf{AQP}(M, q) : & \text{minimize} & x^T(Mx + q) \\ & \text{subject to} & Mx + q \geq 0, \\ & & x \geq 0. \end{array} \tag{1.9}$$

Notice that the objective function of $\mathbf{AQP}(M,q)$ is bounded below by zero if the problem is feasible, it follows immediately from the Frank-Wolfe Theorem (see for example [43]) that $\mathbf{AQP}(M,q)$ admits an optimal solution. Furthermore, the vector $x \in \mathbb{R}^n$ solves the $\mathbf{LCP}(M,q)$ if and only if it is a global minimizer of $\mathbf{AQP}(M,q)$ with a zero objective value. This observation is instrumental for the discovery of several interesting matrix classes.

Motivated by the formulation (1.9), we associate $\mathbf{XLCP}(M,N,\mathcal{P})$ with the following bilinear program [17, 34]:

$$\begin{array}{lll}\mathbf{BLP}(M,N,\mathcal{P}): & \text{minimize} & x^T y \\ & \text{subject to} & Mx - Ny \in \mathcal{P}, \\ & & x \geq 0,\ y \geq 0.\end{array} \tag{1.10}$$

Analogously, the $\mathbf{BLP}(M,N,\mathcal{P})$ plays a similar role in linking properties of the matrix pair $\{M,N\}$ and solution properties of the $\mathbf{XLCP}(M,N,\mathcal{P})$. We shall have more discussions on this in Chapter 2.

Natural Map and Normal Map Formulations

It is not difficult to see that finding a solution to the $\mathbf{LCP}(M,q)$ is equivalent to finding a zero point of the equation

$$g(x) = \min(x, Mx+q), \tag{1.11}$$

where "min(a, b)" denotes the componentwise minimum of two vectors a and b. Also, we notice that $g(x) = 0$ if and only if x satisfies

$$x - h(x) = 0, \tag{1.12}$$

where $h(x) = \max(0, x-(Mx+q))$ and "max(a, b)" denotes the componentwise maximum of two vectors a and b. In other words, x is a fixed-point of h. Recognizing that the max function: $z \mapsto \max(0,z)$ is the Euclidean projector onto the nonnegative orthant cone $\mathbb{R}^n_+$ which we denote $\Pi_{\mathbb{R}^n_+}$, we rewrite (1.12) as

$$x - \Pi_{\mathbb{R}^n_+}(x-(Mx+q)) = 0.$$

This observation can be extended in a more general setting with the help of the Euclidean projector onto a general closed convex set. Before we proceed, we take a pause to formally define Euclidean projector and mention some of its elementary properties.

Definition 1.2.1. Let $\mathbb{E}$ be a finite dimensional Euclidean space with inner product $\langle\cdot,\cdot\rangle$, and let $\mathcal{C} \subseteq \mathbb{E}$ be a closed convex subset, the Euclidean projection of $x \in \mathbb{E}$ onto $\mathcal{C}$ is given by

$$\Pi_{\mathcal{C}}(x) \triangleq \operatorname{argmin}_{y\in\mathcal{C}} \|y-x\|^2. \tag{1.13}$$

It follows from the definition that $\Pi_{\mathcal{C}}(x)$ is the unique minimizer y^* of the strongly convex optimization problem

$$\min_{y\in\mathcal{C}} f(y) \triangleq \|y-x\|^2,$$

which can be characterized by the optimality condition

$$\langle \nabla f(y^*), y-y^*\rangle \geq 0 \quad \forall y \in \mathcal{C}.$$

Noting that $\nabla f(y^*) = 2(y^*-x)$, we have that

$$\langle y - \Pi_{\mathcal{C}}(x), x - \Pi_{\mathcal{C}}(x)\rangle \leq 0 \quad \forall y \in \mathcal{C}. \tag{1.14}$$

If, in addition, $\mathcal{C}$ is a cone, we have the Moreau decomposition [40]: Any $x \in \mathbb{E}$ can be written as

$$x = \Pi_{\mathcal{C}}(x) - \Pi_{\mathcal{C}^*}(-x) \quad \text{with } \langle \Pi_{\mathcal{C}}, \Pi_{\mathcal{C}^*}(-x)\rangle = 0. \tag{1.15}$$

Moreover, the decomposition is unique in the sense that $x = x_1 - x_2$ with $x_1 \in \mathcal{C}, x_2 \in \mathcal{C}^*$ and $\langle x_1, x_2\rangle = 0$ if and only if $x_1 = \Pi_{\mathcal{C}}$ and $x_2 = \Pi_{\mathcal{C}^*}(-x)$.

It is well-known that $\Pi_{\mathcal{C}}$ plays a central role in the study of complementarity problems and variational inequalities. The variational inequality problem is to find a vector $x \in \mathcal{C}$ such that

$$\mathbf{VI}(F,\mathcal{C}): \quad \langle y-x, F(x)\rangle \geq 0 \quad \forall y \in \mathcal{C}. \tag{1.16}$$

It is not difficult to show that if $\mathcal{C}$ is a cone, then $\mathbf{CP}(F,\mathcal{C})$ and $\mathbf{VI}(F,\mathcal{C})$ are in fact equivalent in the sense that a vector x solves (1.7) if and only if it satisfies (1.16) . Because of this close connection, the variational inequality plays an important role in the study of CPs. For example, some important fundamental results such as existence of solutions of CPs are essentially derived from the variational inequality theory. Next, we state the Hartman-Stampacchia's fundamental theorem [24] concerning the existence of solutions for the variational inequality problem. For the completeness, we present the proof using properties of the Euclidean projector and the Brouwer's fixed-point theorem.

Theorem 1.2.1. *Let $\mathbb{E}$ be a finite-dimensional Euclidean space and $\mathcal{C} \subseteq \mathbb{E}$ be an non-empty convex compact set, and let $F : \mathcal{C} \to \mathbb{E}$ be a continuous mapping. Then there exists an $x^* \in \mathcal{C}$ satisfying*

$$\langle x - x^*, F(x^*)\rangle \geq 0 \quad \forall x \in \mathcal{C},$$

*i.e., the **VI**$(F,\mathcal{C})$ has a solution.*

Proof. First, we show that the $\mathbf{VI}(F,\mathcal{C})$ is equivalent to the fixed point equation

$$x = \Pi_{\mathcal{C}}(x - F(x)), \tag{1.17}$$

where $\Pi_{\mathcal{C}}(x)$ denotes the Euclidean projection of x onto the set $\mathcal{C}$. Indeed, if x^* is a fixed-point of (1.17), then it follows that $x^* \in \mathcal{C}$, and by the property of the projection function $\Pi_{\mathcal{C}}$ we deduce

$$\langle x - x^*, (x^* - F(x^*)) - x^*\rangle \leq 0 \quad \forall x \in \mathcal{C},$$

that is, $\langle x - x^*, F(x^*)\rangle \geq 0$ for all $x \in \mathcal{C}$, implying that x^* is a solution of $\mathbf{VI}(F, \mathcal{C})$. Reversing the argument shows that if x^* is a solution of $\mathbf{VI}(F, \mathcal{C})$ then it is also a solution of (1.17). Consider the map

$$H(x) \triangleq \Pi_{\mathcal{C}}(x - F(x)),$$

which is continuous by the continuity of $\Pi_{\mathcal{C}}$. Since $\mathcal{C}$ is a convex compact set, it follow that H admits a fixed-point in $\mathcal{C}$ by the Brouwer's fixed-point theorem, and hence the $\mathbf{VI}(F, \mathcal{C})$ has a solution. □

As a consequence of Theorem 1.2.1, a popular approach for solving the $\mathbf{CP}(F, \mathcal{K})$ is the so-called *natural map equation*

$$F^{\text{nat}}(x) \triangleq x - \Pi_{\mathcal{K}}(x - F(x)) = 0. \tag{1.18}$$

With the help of this observation, we can show that every solution of the CP$(F, \mathcal{K})$ corresponds exactly to a solution to the *normal map equation* (NME)

$$F(\Pi_{\mathcal{K}}(z)) - \Pi_{\mathcal{K}^*}(-z) = 0,$$

which is equivalent to, by the Moreau decomposition (1.15),

$$F^{\text{nor}}(z) \triangleq F(\Pi_{\mathcal{K}}(z)) + z - \Pi_{\mathcal{K}}(z) = 0 \tag{1.19}$$

via $x = \Pi_{\mathcal{K}}(z)$ and $z = x - F(x)$. See, e.g., [11].

Note that the domain of the function F^{nat} is the same as F, whereas the domain of the function F^{nor} is always the entire space $\mathbb{E}$. This feature makes the normal map formulation particularly attractive in algorithm designs in the situation where F is not defined over the entire space $\mathbb{E}$. It should be also pointed out that the main difficulty in solving (1.18) and (1.19) is the non-smoothness of the projector function $\Pi_{\mathcal{K}}$. To overcome this difficulty, in Chapter 4, we develop a smoothing homotopy path for find a solution to the $\mathbf{CP}(F, \mathcal{K})$ based on the normal map reformulation.

1.3 Preliminaries

Euclidean Jordan algebra

In this section, we review concepts in the theory of Euclidean Jordan algebras that are necessary for the purpose of this thesis. See Chapters II–IV of [12] for a more comprehensive discussion on the theory of Euclidean Jordan algebras.

Definition 1.3.1. A *Euclidean Jordan algebra* is a triple $(\mathbb{E}, \circ, \langle\cdot,\cdot\rangle)$, where $\mathbb{E}$ is a finite dimensional vector space over $\mathbb{R}$ equipped with inner product $\langle\cdot,\cdot\rangle$, the Jordan product $(x, y) \mapsto x \circ y : \mathbb{E} \times \mathbb{E} \to \mathbb{E}$ is a bilinear map satisfying:

(i) $x \circ y = y \circ x \quad \forall x, y \in \mathbb{E}$,

(ii) $x \circ (x^2 \circ y) = x^2 \circ (x \circ y) \;\; \forall x, y \in \mathbb{E}$, where $x^2 := x \circ x$,

(iii) $\langle x \circ y, z \rangle = \langle x, y \circ z \rangle \;\; \forall x, y, z \in \mathbb{E}$.

In addition, we assume that there is a unitary element $e \in \mathbb{E}$ such that $x \circ e = x$ for all $x \in \mathbb{E}$.

Definition 1.3.2. Let $\mathbb{E}$ be a Euclidean Jordan algebra. For any $x \in \mathbb{E}$, the *Lyapunov transformation* $L_x : \mathbb{E} \mapsto \mathbb{E}$ is defined as

$$L_x(y) = x \circ y.$$

The *quadratic representation* of x is defined as

$$P_x = 2L_x^2 - L_{x \circ x}.$$

Note that by definition of Euclidean Jordan algebra, L_x, whence P_x, is self-adjoint under $\langle \cdot, \cdot \rangle$. We say that elements x and y operator commute if the Lyapunov transformations L_x and L_y commute, i.e., $L_x L_y = L_y L_x$. It is well known that x and y operator commute if and only if they share the same Jordan frames in their spectral decomposition.

Definition 1.3.3 (Jordan Frame)**.** Let $\mathbb{E}$ be a Euclidean Jordan algebra.

- An element $c \in \mathbb{E}$ is called *idempotent* if $c^2 = c$.
- An idempotent $c \in \mathbb{E}$ is called a *primitive idempotent* if it is nonzero and cannot be written as a sum of two nonzero idempotents.
- Two idempotents c and d are said to be orthogonal if $c \circ d = 0$.
- A *Jordan frame* is a finite set $\{c_1, ..., c_r\}$ of primitive idempotents in $\mathbb{E}$ that are pair-wise orthogonal and sum to the unit e, i.e.,

$$c_i \circ c_j = 0, \text{ if } i \neq j, \text{ and } \sum_{i=1}^{r} c_i = e.$$

  Note that orthogonal idempotents are indeed orthogonal with respect to the inner product $\langle \cdot, \cdot \rangle$ since

$$\langle c_i, c_j \rangle = \langle c_i \circ e, c_j \rangle = \langle e, c_i \circ c_j \rangle = 0.$$

Definition 1.3.4. The *rank* of $\mathbb{E}$ is defined as $r := \max \{m(x) : x \in \mathbb{E}\}$, where $m(x)$ is the degree of $x \in \mathbb{E}$ given by

$$m(x) := \min \left\{ k > 0 : \left\{ e, x, ..., x^k \right\} \text{ is linearly dependent} \right\}.$$

The following theorem says that every element $x \in \mathbb{E}$ can be decomposed in terms of some Jordan frame.

Theorem 1.3.1 (Spectral decomposition, [12]). *Let r be the rank of $\mathbb{E}$. For each $x \in \mathbb{E}$, there exists a Jordan frame $\{c_1, ..., c_r\}$ such that*

$$x = \lambda_1(x)c_1 + \cdots + \lambda_r(x)c_r,$$

where $\lambda_1(x) \geq \cdots \geq \lambda_r(x)$, called the eigenvalues of x, are uniquely determined.

Definition 1.3.5. A closed convex cone $\mathcal{K} \subseteq \mathbb{E}$ is called a *symmetric cone* if it is self-dual, i.e.,

$$\mathcal{K} = \mathcal{K}^* \triangleq \{s \in \mathbb{E} : \langle x, s \rangle \geq 0 \; \forall x \in \mathbb{E}\},$$

and homogeneous, i.e., for any two elements $x, y \in \text{int}(\mathcal{K})$, there exists an invertible linear transformation $\Gamma : \mathbb{E} \to \mathbb{E}$ such that $\Gamma(\mathcal{K}) = \mathcal{K}$ and $\Gamma(x) = y$.

Theorem 1.3.2 (Characterization of Symmetric Cones, [12]). *In a Euclidean Jordan algebra, a symmetric cone is the cone of squares, i.e., $\mathcal{K} = \{x^2 : x \in \mathbb{E}\}$. Moreover, $\mathcal{K}$ coincides with the following equivalent sets:*

(i) the set $\{x \in \mathbb{E} : L_x$ is positive semidefinite under $\langle \cdot, \cdot \rangle\}$;

(ii) the set $\{x \in \mathbb{E} : \lambda_i(x) \geq 0 \; \forall i\}$.

Example 1.3.1. *Three familiar examples of Euclidean Jordan algebras are as follows:*

(i) Euclidean Jordan algebra of n-dimensional vectors:

$$\mathbb{E} = \mathbb{R}^n, \;\; \mathcal{K} = \mathbb{R}^n_+, \;\; r = n, \;\; \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i, \;\; x \circ y = x * y,$$

where $x * y$ denotes the componentwise product of x and y. It is easy to see that the unitary element is $e = (1, 1, ..., 1) \in \mathbb{R}^n$.

(ii) Euclidean Jordan algebra of quadratic forms:

$$\mathbb{E} = \mathbb{R}^n, \;\; \mathcal{K} = \mathcal{L}^n_+ \triangleq \left\{(x_1, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{n-1} : x_1 \geq \|\bar{x}\|\right\}, \;\; r = 2,$$

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i, \;\; x \circ y = (\langle x, y \rangle, x_1 \bar{y} + y_1 \bar{x}).$$

The symmetric cone $\mathcal{L}^n_+$ is called the Lorentz cone or the second order cone. Here, the unitary element is $e = (1, 0, ..., 0) \in \mathbb{R}^n$. Given $x \in \mathbb{E}$ with $\bar{x} \neq 0$, the spectral decomposition of x is given by

$$x = \frac{1}{2}(x_0 + \|\bar{x}\|) \begin{pmatrix} 1 \\ \frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix} + \frac{1}{2}(x_0 - \|\bar{x}\|) \begin{pmatrix} 1 \\ -\frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix}.$$

(iii) Euclidean Jordan algebra of symmetric matrices:

$$\mathbb{E} = \mathcal{S}^n,\ \mathcal{K} = \mathcal{S}^n_+,\ r = n,\ \langle x, y\rangle = tr(xy),\ x \circ y = \frac{1}{2}(xy + yx),$$

where $\mathcal{S}^n$ denotes the set of $n \times n$ real symmetric matrices, $\mathcal{S}^n_+$ denotes the cone of $n \times n$ symmetric positive semidefinite matrices. In this algebra, the unitary element e is the identity matrix I_n. Given any $x \in \mathcal{S}^n$, there exists an orthonormal basis $\{u_1, u_2, ..., u_n\}$ and a set of real numbers $\{\lambda_1, \lambda_2, ..., \lambda_n\}$ satisfying

$$x = \lambda_1 u_1 u_1^T + \cdots + \lambda_n u_n u_u^T.$$

Fix a Jordan frame $\{c_1, ..., c_r\}$ in a Euclidean Jordan algebra $\mathbb{E}$. For $i, j \in \{1, 2, ..., r\}$, defines the eigenspaces

$$V_{ii} \triangleq \{x \in \mathbb{E} : x \circ c_i = x\} = \mathbb{R}c_i$$

and when $i \neq j$,

$$V_{ij} \triangleq \{x \in \mathbb{E} : x \circ c_i = \frac{1}{2}x = x \circ c_j\}.$$

Then we have the following

Theorem 1.3.3 (Peirce decomposition, [12])**.** *Given a Jordan frame $\{c_1, ..., c_r\}$, the space $\mathbb{E}$ decomposes into the orthogonal direct sum of spaces $V_{ij} (i \leq j)$. Furthermore,*

$$\begin{aligned} V_{ij} \circ V_{ij} &\subset V_{ii} + V_{jj}, \\ V_{ij} \circ V_{jk} &\subset V_{ik} \quad \text{if } i \neq k, \\ V_{ij} \circ V_{kl} &= \{0\} \quad \text{if } \{i, j\} \cap \{k, l\} = \emptyset. \end{aligned}$$

Therefore, for any $x \in \mathbb{E}$,

$$x = \sum_{i=1}^{r} x_i c_i + \sum_{1 \leq i < j \leq r} x_{ij},$$

with $x_i \in \mathbb{R}$ and $x_{ij} \in V_{ij}$.

Example 1.3.2. *Let $\mathbb{E} = \mathcal{S}^n$ and consider the canonical Jordan frame $\{E_1, ..., E_n\}$, where E_i is the matrix with one in the (i, i) entry and zero's otherwise. It is easy to verify that*

$$V_{ii} = \{\kappa E_i : \kappa \in \mathbb{R}\} \text{ and } V_{ij} = \{\theta E_{ij} : \theta \in \mathbb{R}\},$$

where E_{ij} is a matrix with one in the (i, j) and (j, i) entries and zero's otherwise. Thus, any $x \in \mathcal{S}^n$ can be written as

$$x = \sum_{i=1}^{n} x_{ii} E_i + \sum_{1 \leq i < j \leq n} x_{ij} E_{ij}.$$

This denotes the Peirce decomposition of x associated with $\{E_1, ..., E_n\}$.

Every Euclidean Jordan algebra can be written as a direct sum of simple ideals (see, e.g., [12]), each of which is a Euclidean Jordan algebra under the induced inner product. This means that every symmetric cone can be written as the orthogonal direct sum of irreducible symmetric cones, whence reducing the classification of symmetric cones to that of irreducible symmetric cones. This decomposition of Euclidean Jordan algebra will be instrumental in the definition of Cartesian P-property in Chapter 3.

Chapter 2

The Weighted Extended Linear Complementarity Problem

2.1 Introduction

Let $M, N \in \mathbb{R}^{m\times n}$ be two real matrices of order $m \times n$, and let $\mathcal{P} \subseteq \mathbb{R}^m$ be a polyhedral and $w \in \mathbb{R}^n_+$ be a given vector. The weighted extended linear complementarity problem (wXLCP) is to find a pair of vectors $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ satisfying

$$\mathbf{wXLCP}(w, M, N, \mathcal{P}): \quad x,\ y \geq 0, \quad Mx - Ny \in \mathcal{P} \quad \text{and} \quad x \circ y = w. \tag{2.1}$$

Examples of $\mathbf{wXLCP}(w, M, N, \mathcal{P})$ unifies several formulations introduced in the first chapter. Specifically,

- When $m = n, w = 0, N = I$ and $\mathcal{P} = \{-q\}$ for some $q \in \mathbb{R}^n$, we obtain the $\mathbf{LCP}(M, q)$ in (1.1);
- When $m = n, w = 0$ and $\mathcal{P} = \{q\}$ for some $q \in \mathbb{R}^n$, we obtain the $\mathbf{HLCP}(M, N, q)$ in (1.4);
- When $w = 0$, we obtain the $\mathbf{XLCP}(M, N, \mathcal{P})$ in (1.5);
- When $m = n$ and $\mathcal{P} = \{q\}$ for some $q \in \mathbb{R}^n$, we obtain the $\mathbf{wHLCP}(w, M, N, q)$ in (1.6).

Motivated by the important roles played by matrix classes in the LCP theory, Gowda [17] described and characterized the *column sufficiency*, the *row sufficiency* and the P-property in the setting of the XLCP, thereby generalizing the corresponding notions and results from LCPs and HLCPs. The main purpose of this chapter is to undertake a further study of several matrix-pair theoretic properties and investigate their relationships to solution properties of wXLCPs.

Notations and Terminalogy

Given a non-empty polyhedral $\mathcal{P} = \{u \in \mathbb{R}^m : Lu \geq b\}$ for some matrix $L \in \mathbb{R}^{k\times m}$ and vector $b \in \mathbb{R}^k$, the recession cone is defined as

$$0^+\mathcal{P} \triangleq \{u \in \mathbb{R}^m : Lu \geq 0\}.$$

The dual of the recession cone can be written as

$$(0^+\mathcal{P})^* = \{L^T\lambda : \ \lambda \geq 0,\ \lambda \in \mathbb{R}^k\}.$$

The feasible set of $\mathbf{wXLCP}(w, M, N, \mathcal{P})$ is defined to be the set

$$\text{FEA}(M, N, \mathcal{P}) \triangleq \{(x, y) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+ : Mx - Ny \in \mathcal{P}\},$$

and we say that $\mathbf{wXLCP}(M, N, \mathcal{P})$ is *feasible* if $\text{FEA}(M, N, \mathcal{P}) \neq \emptyset$. A pair $(x, y) \in \text{FEA}(M, N, \mathcal{P})$ is called a *strictly feasible* point if $x > 0$ and $y > 0$. Note the the feasible set is independent of the weight vector w. We denote the set of solutions of $\mathbf{wXLCP}(w, M, N, \mathcal{P})$ as

$$\text{SOL}(w, M, N, \mathcal{P}) \triangleq \{(x, y) \in \text{FEA}(M, N, \mathcal{P}) : \ x \circ y = w\}.$$

For notational simplicity, we denote by u^2 and $\sqrt{u}$ the vector whose i-th component is equal to u_i^2 and $\sqrt{u_i}$, respectively. For two vectors u, v of the same dimension, we denote by $\frac{u}{v}$ the vector whose i-th component is defined as

$$(\frac{u}{v})_i \triangleq \begin{cases} u_i/v_i & \text{if } u_iv_i \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Given matrices $M, N \in \mathbb{R}^{m\times n}$, we say that $\{\overline{M}, \overline{N}\}$ is a *column rearrangement* of $\{M, N\}$ if for each index i, $\overline{M}_i = M_i$ and $\overline{N}_i = N_i$, or $\overline{M}_i = N_i$ and $\overline{N}_i = M_i$ where the subscript refers to the corresponding column. We also say that a pair $\{M', N'\}$ is a *twisted column rearrangement* of $\{M, N\}$ if for each index i, $M'_i = M_i$ and $N'_i = N_i$, or $M'_i = -N_i$ and $N'_i = -M_i$. It is not difficult to see that if (x, y) is a solution of $\mathbf{wXLCP}(w, M, N, \mathcal{P})$, then a suitable permutation of components in (x, y) will yield a solution of $\mathbf{wXLCP}(w, M', N', \mathcal{P})$.

We say that a vector is a KKT vector of a constrained optimization problem if the vector together with some dual vectors satisfy the KKT conditions.

2.2 Optimization Formulations

In this section we introduce two optimization formulations for the $\mathbf{wXLCP}(w, M, N, \mathcal{P})$. Throughout this section, we assume without loss of generality that $\mathcal{P} = \{u \in \mathbb{R}^m : Lu \geq b\}$ with $L \in \mathbb{R}^{k\times m}$ and $b \in \mathbb{R}^k$.

The constrained optimization formulation

Associated with the $\mathbf{wXLCP}(w, M, N, \mathcal{P})$ is the following constrained optimization problem on the same feasible region:

$$\begin{aligned} \mathbf{wBLP}(w, M, N, \mathcal{P}): \quad & \text{minimize} \quad x^T y - \sum_{i=1}^{n} w_i \log x_i y_i \\ & \text{subject to} \quad Mx - Ny \in \mathcal{P}, \\ & \qquad\qquad\quad x \geq 0,\ y \geq 0. \end{aligned} \tag{2.2}$$

Clearly, (2.2) is a direct extension of the $\mathbf{BLP}(M, N, \mathcal{P})$ in (1.10). It should be noted that the $\mathbf{BLP}(M, N, \mathcal{P})$ always has an optimal solution if it is feasible, whereas $\mathbf{wBLP}(w, M, N, \mathcal{P})$ does not necessarily have this nice property as the Frank-Wolfe theorem is not applicable. For convenience of discussions, we denote the index sets

$$\overline{\mathcal{W}} \triangleq \{i \in [n] : w_i > 0\} \quad \text{and} \quad \widehat{\mathcal{W}} \triangleq \{i \in [n] : w_i = 0\}.$$

For any vector $u \in \mathbb{R}^n$, we denote by $u_{\overline{\mathcal{W}}}$ and $u_{\widehat{\mathcal{W}}}$, respectively, the vector formed by components of u that correspond to indices in $\overline{\mathcal{W}}$ and $\widehat{\mathcal{W}}$, respectively. We note that, if (x, y) is in the domain of the objective function of $\mathbf{wBLP}(w, M, N, \mathcal{P})$ then we have $x_{\overline{\mathcal{W}}} y_{\overline{\mathcal{W}}} > 0$.

Let (x, y) satisfy the KKT conditions for (2.2), then there exist vectors $\lambda \in \mathbb{R}^k_+, \alpha, \beta \in \mathbb{R}^n_+$ such that

$$y - \frac{w}{x} - \alpha - M^T L^T \lambda = 0, \tag{2.3}$$

$$x - \frac{w}{y} - \beta + N^T L^T \lambda = 0, \tag{2.4}$$

and

$$\alpha^T x = \beta^T y = \lambda^T (LMx - LNy - b) = 0. \tag{2.5}$$

In next section we shall study necessary and sufficient conditions for a KKT vector of the $\mathbf{wBLP}(w, M, N, \mathcal{P})$ to be a solution of the $\mathbf{wXLCP}(w, M, N, \mathcal{P})$.

The unconstrained optimization formulation

Inspired by the work in [36, 37, 51, 26] where HLCPs, XLCPs and linearly constrained convex problems are reformulated as optimization problems with simple bound constraints or even unconstrained minimization problems, we reformulate the $\mathbf{wXLCP}(w, M, N, \mathcal{P})$ as an unconstrained optimization problem of the form

$$\min_{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n} F(x, y) \triangleq P(x, y) + \Phi(x, y)$$

where

$$P(x, y) = \begin{cases} 0, & \text{if } LMx - LNy \geq b \\ > 0, & \text{otherwise} \end{cases}$$

and

$$\Phi(x,y)=0 \iff x,y\geq 0 \text{ and } x\circ y=w.$$

A possible choice of $P(x,y)$ is the exterior penalty function for the polyhedral $\{(x,y)\in\mathbb{R}^n\times\mathbb{R}^n : LMx-LNy\geq b\}$, i.e.,

$$P(x,y)=\|\Pi_{\mathbb{R}^k_+}(-LMx+LNy+b)\|^2.$$

Recall that $\Pi_{\mathbb{R}^k_+}(z)$ denotes the Euclidean projector of $z\in\mathbb{R}^k$ onto $\mathbb{R}^k_+$. Note that $P(x,y)$ is a continuously differentiable convex function.

Besides the natural and normal map reformulations, another popular way to solve complementarity problems is to reformulate them as a system of nonlinear equations via NCP-functions. The key property of an NCP-function $\varphi:\mathbb{R}^2\to\mathbb{R}$ is as follows:

$$\varphi(a,b)=0 \iff a\geq 0,\ b\geq 0 \text{ and } ab=0.$$

The article [52] and references therein contain an excellent review on various NCP-functions and their properties. Among many NCP-functions the Fischer-Burmeister function [15]

$$\varphi_{FB}(a,b)=\sqrt{a^2+b^2}-a-b$$

has been extensively studied. In particular, it is known that φ_{FB} is differentiable at any $(a,b)\neq(0,0)$ and φ_{FB}^2 is continuously differentiable everywhere.

For a fixed $w\geq 0$, we consider a modification of φ_{FB} defined by

$$\phi(a,b)=(\sqrt{a^2+b^2+2w}-a-b)^2. \tag{2.6}$$

The following lemma summarizes a few useful properties of ϕ.

Lemma 2.2.1. *Let $w\geq 0$ be fixed and ϕ be defined in* (2.6)*. The following statements hold:*

(i) $\phi(a,b)=0 \iff a\geq 0,\ b\geq 0,\ ab=w.$

(ii) *ϕ is continuously differentiable at every point $(a,b)\in\mathbb{R}^2$.*

(iii) $\frac{\partial\phi}{\partial a}(a,b)\frac{\partial\phi}{\partial b}(a,b)\geq 0\ \forall(a,b)\in\mathbb{R}^2.$

(iv) $\frac{\partial\phi}{\partial a}(a,b)\frac{\partial\phi}{\partial b}(a,b)=0 \implies \phi(a,b)=0.$

Proof. The claims can be verified easily, and the proof follows from Lemma 2.1 in [26]. In particular, $\frac{\partial\phi}{\partial a}(0,0)=\frac{\partial\phi}{\partial b}(0,0)=0$. □

The above lemma suggests a choice of $\Phi(x,y)$ as follows

$$\Phi(x,y)=\|\sqrt{x^2+y^2+2w}-x-y\|^2 \triangleq \sum_{i=1}^n\phi(x_i,y_i).$$

Therefore, we can reformulate (2.2) as the following equivalent unconstrained optimization problem:

$$\min_{(x,y)\in\mathbb{R}^n\times\mathbb{R}^n} \quad F(x,y) \triangleq \|(-LMx+Ny+q)_+\|^2 + \Phi(x,y). \tag{2.7}$$

By the differentiability of P and Φ, the objective function F is continuously differentiable. Clearly, if (x, y) solves (2.7) with optimal value of zero, then it is a solution of (2.2), and vice versa. Unfortunately, F is not convex in general, so its stationary points are not necessarily global optimal. In next section, we characterize conditions under which the stationary point of (2.7) is necessarily a solution of the $\mathbf{wXLCP}(w, M, N, \mathcal{P})$.

2.3 The Monotonicity

One of the most well-known notions in the field of mathematical programming is the monotonicity. Recall that $M \in \mathbb{R}^{n\times n}$ (not nessesarily symmetric) is called a monotone matrix if $x^T Mx \geq 0$ for all $x \in \mathbb{R}^n$. Under the monotonicity of M, the $\mathbf{LCP}(M, q)$ has a polyhedral solution set and its feasibility implies the solvability. Moreover, the quadratic programming reformulation (1.9) is convex and any of its KKT vector is a solution to the $\mathbf{LCP}(M, q)$.

We describe the monotonicity for a pair of non-square matrices, and extend corresponding results in the LCP to the setting of wXLCP.

Definition 2.3.1. Given matrices $M, N \in \mathbb{R}^{m\times n}$ and a polyhedral $\mathcal{P} \subseteq \mathbb{R}^m$. We say that the pair $\{M, N\}$ has the *monotonicity* property with respect to $\mathcal{P}$ if the following condition holds:

$$Mx - Ny \in \mathcal{P} - \mathcal{P} \Longrightarrow \langle x, y\rangle \geq 0. \tag{2.8}$$

When $m = n$ and $\mathcal{P} = \{q\}$ for some $q \in \mathbb{R}^m$, the condition (2.8) reduces to the column monotonicity defined in [58]. It was proved there that the column monotonicity of $\{M, N\}$ is equivalent to: (i) $M + N$ is nonsingular, and (ii) MN^T is monotone. The following example shows that the squareness of M and N is necessary for the monotonicity of MN^T. Take the pair

$$A = \begin{pmatrix} 3/2 & -1/2 \\ 1/2 & -1/2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix},$$

which is column monotone [58], and let $M = (A; B)$ and $N = (B; A)$. Direct verification shows that $\{M, N\}$ satisfies (2.8) but the matrix MN^T is not monotone.

The proposition below describes several properties of the solution set of a wXLCP under the monotonicity property.

Proposition 2.3.1. *Let $M, N \in \mathbb{R}^{m\times n}$, $w \in \mathbb{R}^n_+$ and $\mathcal{P} \subseteq \mathbb{R}^m$ be a polyhedral. The function*

$$f(x,y) = x^T y - \sum_{i=1}^n w_i \log x_i y_i$$

is convex on FEA$(M, N, \mathcal{P})$ if $\{M, N\}$ is monotone respect to $\mathcal{P}$.

Proof. We can write $f(x,y) = h(x,y) + g(x,y)$ with

$$h(x,y) = x^T y \quad \text{and} \quad g(x,y) = \sum_{i=1}^{n} (-w_i \log x_i - w_i \log y_i).$$

Note that $g(x,y)$ is convex as each individual term in the summation is convex. Thus it suffices to show that $h(x,y)$ is convex on FEA$(M,N,\mathcal{P})$. Let $(\bar{x},\bar{y}),(\hat{x},\hat{y}) \in$ FEA$(M,N,\mathcal{P})$ and $\lambda \in [0,1]$, easy calculations yield that

$$\begin{aligned} & \lambda h(\bar{x},\bar{y}) + (1-\lambda)h(\hat{x},\hat{y}) - h(\lambda\bar{x} + (1-\lambda)\hat{x}, \lambda\bar{y} + (1-\lambda)\hat{y}) \\ = \; & \lambda(1-\lambda)(\bar{x}-\hat{x})^T(\bar{y}-\hat{y}). \end{aligned}$$

Since $(\bar{x}-\hat{x}, \bar{y}-\hat{y})$ satisfies the left-hand side condition of (2.8), the desired claim follows immediately. □

We state the implications of Proposition 2.3.1 as a corollary.

Corollary 2.3.1. *Let $M,N \in \mathbb{R}^{m\times n}$, $w \in \mathbb{R}^n_+$ and $\mathcal{P} \subseteq \mathbb{R}^m$ be a polyhedral. If $\{M,N\}$ is monotone with respect to $\mathcal{P}$, then the following statements hold:*

(i) SOL$(w,M,N,\mathcal{P})$ is convex (possibly empty);

*(ii) any KKT solution of (2.2), if exists, is a solution of **wXLCP**$(w,M,N,\mathcal{P})$;*

(iii) if $w = 0$ and FEA$(w,M,N,\mathcal{P}) \neq \emptyset$, then SOL$(w,M,N,\mathcal{P})$ is a non-empty polyhedron.

Proof. Part (i) and (ii) follow from the fact that (2.2) is a convex optimization problem over a polyhedral. Part (iii) follows from Theorem 3.1 in [34]. In particular, when $w = 0$, (2.2) reduces to a quadratic program over a polyhedron, and the existence of solutions is thus implied by the Frank-Wolfe theorem. Since the set of optimal solutions of any quadratic program is equal to the union of a finite number of convex polyhedra, the convexity of SOL$(w,M,N,\mathcal{P})$ implies its polyhedrality. □

2.4 The Column Sufficiency

A matrix $M \in \mathbb{R}^{n\times n}$ is said to have the *column sufficiency* property if the condition

$$x \circ Mx \le 0 \implies x \circ Mx = 0$$

holds for any $x \in \mathbb{R}^n$. This property characterizes the convexity of the solution set of the **LCP**(M,q) for any $q \in \mathbb{R}^n$. In [17], this notion was formulated in the context of the XLCP.

Definition 2.4.1. Given matrices $M, N \in \mathbb{R}^{m\times n}$ and a polyhedral $\mathcal{P} \subseteq \mathbb{R}^m$, we say that $\{M, N\}$ has the *column sufficiency property* with respect to $\mathcal{P}$ if the following condition holds

$$\left.\begin{aligned} Mx - Ny \in \mathcal{P} - \mathcal{P} \\ x \circ y \leq 0 \end{aligned}\right\} \implies x \circ y = 0. \tag{2.9}$$

Note that this property is invariant under (twisted) column rearrangements of $\{M, N\}$, and it implies that $M + N$ has a full rank. Thus, when $m = n$, there is a (twisted) column rearrangement $\{\overline{M}, \overline{N}\}$ of $\{M, N\}$ such that $\overline{M}$ is nonsingular, and hence we can rewrite (2.9) so that one of the matrices involved in the reformulation is the identity matrix. These ideas have also been discussed in [18] for the purpose of reducing a monotone HLCP to an equivalent LCP.

Parallel to the result that M is sufficient matrix if and only if the $\mathbf{LCP}(M, q)$ has a convex solution set, Gowda [17] established the equivalence between the column sufficiency of $\{M, N\}$ and the convexity of the solution set of $\mathbf{XLCP}(M, N, \mathcal{P})$.

Theorem 2.4.1 ([17])**.** *The pair $\{M, N\}$ has the column sufficiency property with respect to $\mathcal{P}$ if and only if the **XLCP**$(M, N, \mathcal{P} + p)$ has a convex solution set for each $p \in \mathbb{R}^m$.*

We show that this result can be further extended in the context of the wXLCP.

Proposition 2.4.1. *The pair $\{M, N\}$ has the column sufficiency property with respect to $\mathcal{P}$ if and only if for each $p \in \mathbb{R}^m$ and each $w \in \mathbb{R}^n_+$, SOL$(w, M, N, \mathcal{P} + p)$ is convex.*

Proof. Let $w \in \mathbb{R}^n_+$ and $p \in \mathbb{R}^m$ be arbitrary. If $\mathrm{SOL}(w, M, N, \mathcal{P} + p)$ contains at most one element, then there is nothing to prove. Let $(\bar{x}, \bar{y}), (\hat{x}, \hat{y}) \in \mathrm{SOL}(w, M, N, \mathcal{P} + p)$ be two distinct solutions. For each $i \in \overline{\mathcal{W}}$, we have $\bar{x}_i\bar{y}_i = \hat{x}_i\hat{y}_i = w_i > 0$, it follows that $\bar{x}_i \geq \hat{x}_i$ if and only if $\bar{y}_i \leq \hat{y}_i$, i.e., $(\bar{x} - \hat{x})_i(\bar{y} - \hat{y})_i \leq 0$. If $i \in \widehat{\mathcal{W}}$ then $\bar{x}_i\bar{y}_i = \hat{x}_i\hat{y}_i = 0$, we deduce that $(\bar{x} - \hat{x})_i(\bar{y} - \hat{y})_i = -\bar{x}_i\hat{y}_i - \hat{x}_i\bar{y}_i \leq 0$. Thus, the pair $(u, v) = (\bar{x} - \hat{x}, \bar{y} - \hat{y})$ satisfies the left-hand side condition in (2.9), implying that $u \circ v = 0$ by the column sufficiency of $\{M, N\}$. Consequently, we obtain

$$2w - \bar{x} \circ \hat{y} - \hat{x} \circ \bar{y} = 0.$$

Let $t \in [0, 1]$, then easy calculations yield

$$[t\bar{x} + (1 - t)\hat{x}] \circ [t\bar{y} + (1 - t)\hat{y}] = w.$$

Conversely, the hypothesis implies that the solution set of the $\mathbf{XLCP}(M, N, \mathcal{P} + p)$ is convex for any p. Consequently, the column sufficiency of {M, N} follows from Theorem (2.4.1). This completes the proof. □

The following example from [44] shows that the convexity of $\mathrm{SOL}(w, M, N, \mathcal{P})$ for some nonzero w does not necessarily imply the column sufficiency of $\{M, N\}$.

Example 2.4.1. *Let* $m = n = 2, \mathcal{P} = \{(-3,0)^T\}$ *and* $w = (2,0)^T$, *and let*

$$M = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \quad N = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

It is easy to verify that SOL$(w, M, N, \mathcal{P})$ *is given by the following convex set*

$$\{(x,y) \in \mathbb{R}^2_+ \times \mathbb{R}^2_+ : x_1 = 1, y_1 = 2, y_2 = 0\}$$

However, if we take $u = -v = (-1,0)^T$, *the left-hand side condition in* (2.9) *is satisfied but* $u_1 v_1 = -1 < 0$. *This shows that the pair* $\{M, N\}$ *is not column sufficient.*

2.5 The Row Sufficiency

Another important property in the theory of the LCP is the so-called *row sufficiency*. A matrix $M \in \mathbb{R}^{n\times n}$ is said to have the row sufficiency if M^T has the column sufficiency, i.e., $x \circ M^T x \leq 0 \Longrightarrow x \circ M^T x = 0$. A property characterizing the row sufficiency is that for every $q \in \mathbb{R}^n$, every KKT vector of (1.9) solves the **LCP**(M, q). This notion was extended to the HLCP and a similar characterization was established in [34]. Later on, Gowda [17] formulated the concept in the context of the XLCP.

Definition 2.5.1. Given matrices $M, N \in \mathbb{R}^{m\times n}$ and a polyhedral $\mathcal{P} \subseteq \mathbb{R}^m$, we say that $\{M, N\}$ has the *row sufficiency property* with respect to $\mathcal{P}$ if the following implication holds

$$\left.\begin{array}{r} (M^T u) \circ (N^T u) \leq 0 \\ u \in (0^+\mathcal{P})^* \end{array}\right\} \Longrightarrow (M^T u) \circ (N^T u) = 0. \tag{2.10}$$

Note that this property is invariant under both column rearrangements and twisted column rearrangements of $\{M, N\}$.

Mangasarian and Pang [34] showed that when MN^T is copositive over $(0^+\mathcal{P})^*$, that is,

$$\langle u, MN^T u\rangle \geq 0 \quad \forall u \in (0^+\mathcal{P})^*,$$

then every KKT vector of (1.10) solves the **XLCP**$(M, N, \mathcal{P})$. Gowda [17] established how a row sufficient matrix-pair characterises the relationship between a KKT vector of the program (1.10) and a solution to the **XLCP**$(M, N, \mathcal{P})$.

Theorem 2.5.1 ([17]). *The pair* $\{M, N\}$ *is row sufficient with respect to* $\mathcal{P}$ *if and only if for every* $p \in \mathbb{R}^m$, *and for every column rearrangement* $\{\overline{M}, \overline{N}\}$ *of* $\{M, N\}$, *every KKT vector of*

$$\begin{array}{ll} \textit{minimize} & x^T y \\ \textit{subject to} & \overline{M}x - \overline{N}y \in \mathcal{P} + p \\ & x \geq 0, \ y \geq 0. \end{array} \tag{2.11}$$

is a solution of $\textbf{\textit{XLCP}}(\overline{M}, \overline{N}, \mathcal{P} + p)$.

With the help of the program (2.2), we prove the following results.

Proposition 2.5.1. *Let $M, N \in \mathbb{R}^{m\times n}$ and $\mathcal{P} \subseteq \mathbb{R}^m$ be a polyhedral. The following statements hold:*

(a) The pair $\{M, N\}$ is row sufficient with respect to $\mathcal{P}$ if and only if for every $p \in \mathbb{R}^m$, and for every column rearrangement $\{\overline{M}, \overline{N}\}$ of $\{M, N\}$ and for every $w \in \mathbb{R}^n_+$, every KKT vector of

$$\begin{aligned} &\textit{minimize} \quad x^T y - \sum_{i=1}^{n} w_i \log x_i y_i \\ &\textit{subject to} \quad \overline{M}x - \overline{N}y \in \mathcal{P} + p \\ &\qquad\qquad\quad x \geq 0,\ y \geq 0. \end{aligned} \tag{2.12}$$

is a solution of $\boldsymbol{wXLCP}(w, \overline{M}, \overline{N}, \mathcal{P} + p)$.

(b) If there exists $w \in \mathbb{R}^n_+$ such that for every $p \in \mathbb{R}^m$ and every column rearrangement $\{\overline{M}, \overline{N}\}$ of $\{M, N\}$, every KKT vector of (2.12) is a solution of $\boldsymbol{wXLCP}(w, \overline{M}, \overline{N}, \mathcal{P}+p)$ *and, in addition, the following implication holds*

$$(\overline{M}^T u)_i \iff (\overline{N}^T u)_i = 0 \quad \forall u \in (0^+\mathcal{P})^*,\ \forall i \in [n],$$

then the pair $\{M, N\}$ has the row sufficiency property with respect to $\mathcal{P}$.

Proof. Without loss of generality, we assume that the polyhedral $\mathcal{P}$ is given by

$$\mathcal{P} = \{u \in \mathbb{R}^m : Lu \geq b\}$$

where $L \in \mathbb{R}^{k\times m}$ and $b \in \mathbb{R}^m$.

(a). Since the row sufficiency property of $\{M, N\}$ is invariant under a translation of the polyhedral $\mathcal{P}$, it suffices to show the "only if" part for the **wXLCP**$(w, M, N, \mathcal{P})$. Let (x, y) be a KKT vector of the optimization problem

$$\begin{aligned} &\text{minimize} \quad x^T y - \sum_{i=1}^{n} w_i \log x_i y_i \\ &\text{subject to} \quad LMx - LNy \geq b \\ &\qquad\qquad\quad x \geq 0,\ y \geq 0. \end{aligned} \tag{2.13}$$

Then there exist vectors $\lambda \in \mathbb{R}^k_+, \alpha, \beta \in \mathbb{R}^n_+$ such that

$$y - \frac{w}{x} - \alpha - M^T L^T \lambda = 0,$$

$$x - \frac{w}{y} - \beta + N^T L^T \lambda = 0,$$

and

$$\alpha^T x = \beta^T y = \lambda^T(LMx - LNy - b) = 0.$$

Note that for each $i \in \overline{\mathcal{W}}$, we have $x_i y_i > 0$ and thus

$$\begin{aligned} & (M^T L^T \lambda)_i (N^T L^T \lambda)_i \\ = & \ (y_i - w_i/x_i - \alpha_i)(-x_i + w_i/y_i + \beta_i) \\ = & \ -\frac{(x_i y_i - w_i)^2}{x_i y_i} \le 0. \end{aligned} \tag{2.14}$$

For $i \in \widehat{\mathcal{W}}$, $(M^T L^T \lambda)_i (N^T L^T \lambda)_i$ reduces to

$$(y_i - \alpha_i)(-x_i + \beta_i) = -x_i y_i - \alpha_i \beta_i \le 0. \tag{2.15}$$

Let $u = L^T \lambda$, then $u \in (0^+\mathcal{P})^*$ and $(M^T u) \circ (N^T u) \le 0$. By the row sufficiency of $\{M, N\}$, it follows that $(M^T u) \circ (N^T u) = 0$. In view of (2.14) and (2.15), we conclude that $x \circ y = w$, showing that (x, y) is a solution of the **wXLCP**$(w, M, N, \mathcal{P})$. Clearly, the "if" part follows from Theorem 2.5.1.

(b). Let $w \in \mathbb{R}^n_+$ be given. We shall show that if the row sufficiency of $\{M, N\}$ is violated, then there exists some appropriate $p^* \in \mathbb{R}^m$ such that a KKT vector (x^*, y^*) of the problem

$$\begin{aligned} \text{minimize} \quad & x^T y - \sum_{i=1}^n w_i \log x_i y_i \\ \text{subject to} \quad & LMx - LNy \ge b + Lp^* \\ & x \ge 0, \ y \ge 0. \end{aligned} \tag{2.16}$$

is not a solution of the **wXLCP**$(w, M, N, \mathcal{P} + p^*)$.

Assume that $\{M, N\}$ is not row sufficient, then there is a vector $u^* \in (0^+\mathcal{P})^*$ satisfying

$$(M^T u^*) \circ (N^T u^*) \le 0 \quad \text{and} \quad (M^T u^*)_j (N^T u^*)_j < 0 \text{ for some } j.$$

By interchanging columns of M_i and N_i and working with a column rearrangement of $\{M, N\}$ if necessary, we may assume without loss of generality that $(M^T u^*)_j > 0$ and $(N^T u^*)_j < 0$. For the ease of notation, we denote

$$\mu = M^T u^* \quad \text{and} \quad \nu = N^T u^*.$$

Defining, for each $i \in \widehat{\mathcal{W}}$, i.e., $w_i = 0$,

$$x_i^* = \max\{-\nu_i, 0\}, \quad y_i^* = \max\{\mu_i, 0\},$$

$$\alpha_i^* = \max\{-\mu_i, 0\}, \quad \beta_i^* = \max\{\nu_i, 0\}.$$

Since $\mu \circ \nu \le 0$, we have

$$\alpha_i^* x_i^* = \beta_i^* y_i^* = 0.$$

Also, by definition, it holds that

$$y_i^* - \alpha_i^* - \mu_i = 0, \quad x_i^* - \beta_i^* + \nu_i = 0. \tag{2.17}$$

We next determine x_i^*, y_i^*, α_i^* and β_i^* for $i \in \overline{\mathcal{W}}$ satisfying the equations

$$y_i - w_i/x_i - \alpha_i - \mu_i = 0, \tag{2.18}$$

$$x_i - w_i/y_i - \beta_i + \nu_i = 0. \tag{2.19}$$

Since $w_i > 0$ for $i \in \overline{\mathcal{W}}$, we have the additional constraints $x_i^* y_i^* > 0$ and $\alpha_i^* = \beta_i^* = 0$. By the assumption on the pair $\{M, N\}$, we have the following three possibilities:

(i) If $\mu_i > 0, \nu_i < 0$, we solve equations (2.18) and (2.19)

$$x_i^* = w_i/y_i^* - \nu_i, \quad y_i^* = w_i/x_i^* + \mu_i \tag{2.20}$$

and obtain

$$x_i^* = \frac{-\mu_i\nu_i + \sqrt{\mu_i^2\nu_i^2 - 4\mu_i\nu_i w_i}}{2\mu_i} > 0, \quad y_i^* = w_i/x_i^* + \mu_i > 0.$$

(ii) Similarly, if $\mu_i < 0, \nu_i > 0$, we set

$$y_i^* = \frac{\mu_i\nu_i + \sqrt{\mu_i^2\nu_i^2 - 4\mu_i\nu_i w_i}}{2\nu_i} > 0, \quad x_i^* = \frac{w_i}{y_i^* - \mu_i} > 0.$$

(iii) If $\mu_i = \nu_i = 0$, set $x_i^* = w_i$ and $y_i^* = 1$.

In summary, by (2.17), (2.18) and (2.19), we have determined nonnegative vectors x^*, y^*, α^* and β^* such that

$$y^* - \frac{w}{x^*} - \alpha^* - (M^T u^*) = 0, \quad x^* - \frac{w_i}{y^*} - \beta^* + (N^T u^*) = 0. \tag{2.21}$$

and

$$(x^*)^T \alpha^* = (y^*)^T \beta^* = 0. \tag{2.22}$$

It now comes to the construction of the vector $p^* \in \mathbb{R}^m$. Let $c := L(Mx^* - Ny^*) - b$. It can be easily verified that

$$\{p \in \mathbb{R}^m : Lp \leq c\} = (Mx^* - Ny^*) - \mathcal{P},$$

and the latter is clearly nonempty. Note that the set $\{\lambda \in \mathbb{R}^k : L^T\lambda = u^*, \lambda \geq 0\}$ is also nonempty since $u^* \in (0^+\mathcal{P})^*$. Consider the linear programming

$$\min_{\lambda}\{\lambda^T c : \ L^T\lambda = u^*, \lambda \geq 0\}$$

and its dual

$$\max_p \{p^T u^* : \ Lp \leq c\}.$$

Since both of them are feasible, it follows from the strong duality that they admit optimal solutions, say λ^* and p^*, respectively, and $(\lambda^*)^T c = (p^*)^T u^*$. Now the inequality $Lp^* \leq c$ implies that the pair (x^*, y^*) constructed above is a feasible point of the problem (2.16). Furthermore,

$$(\lambda^*)^T (LMx^* - LNy^* - Lp^* - b) = (\lambda^*)^T (c - Lp^*) = (\lambda^*)^T c - (p^*)^T u^* = 0,$$

which, together with (2.21) and (2.22), shows that (x^*, y^*) is a KKT vector of (2.16) with corresponding dual variables α^*, β^* and λ^*. Nevertheless, $(x^*, y^*) \notin \mathrm{SOL}(w, M, N, \mathcal{P})$. Indeed, if $j \in \widehat{\mathcal{W}}$, then $x_j^* y_j^* = -\mu_j \nu_j > 0$. If $j \in \overline{\mathcal{W}}$, for instance $\mu_j > 0, \nu_j < 0$, then it follows from (2.20) that $x_j^* y_j^* = \mu_j x_j + w_j > w_j$. $\square$

A remarkable result in the XLCP (and hence the LCP and the HLCP) theory is that, under the row sufficiency property, the feasibility of the problem implies its solvability. This implication no longer holds in the setting of wXLCPs. It is not difficult to see that if the wXLCP has no strict feasible point but the weight vector $w \in \mathbb{R}^n_{++}$, then clearly no solution exists. For a concrete example, we take

$$M = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad N = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Straightforward calculations show that the pair $\{M, N\}$ is row sufficient with respect to $\mathcal{P} = \{(q, 0)^T\}$ for any $q \in \mathbb{R}$. The feasible set can be expressed as

$$\{(x, y) \in \mathbb{R}^2_+ \times \mathbb{R}^2_+ : \ x_1 + x_2 - y_1 = q, \ y_2 = 0\},$$

which is clearly nonempty. However, if the weight vector has $w_2 > 0$, then it is impossible to have $x_2 y_2 = w_2$ since $y_2 = 0$ for any feasible point.

We now seek to derive more solution properties of $\mathbf{wXLCP}(w, M, N, \mathcal{P})$ with the help of the unconstrained optimization problem (2.7). There are two main advantages of reformulating problems as ones with simple box contraints or even no contraint. Firstly, the resulting problem significantly reduces the number of constraints especially when it is large. Secondly, it allows us to apply many efficient algorithms for solving box constrained and uncontrained problems even if the number of variables is large. Nevertheless, the reformulation is not guaranteed to be convex, hence its stationary point is not necessarily a global minimizer. This motivates us to characterize conditions under which the stationary point of (2.7) is necessarily a solution of $\mathbf{wXLCP}(w, M, N, \mathcal{P})$. The desired property enables us to apply local methods guaranteed to converge to a local optimum for solving wXLCPs.

The following result gives some critical properties of stationary points of (2.7) which are not solutions of (2.2).

Theorem 2.5.2. *Assume that* $FEA(w, M, N, \mathcal{P}) \neq \emptyset$. *Let* (x^*, y^*) *be a stationary point of* (2.7) *and* $u^* = L^T \Pi_{\mathbb{R}^m_+}(-Mx^* + Ny^* + q)$. *If* $(x^*, y^*) \notin SOL(w, M, N, \mathcal{P})$, *then the following statements hold:*

(i) $\Phi(x^*, y^*) > 0$;

(ii) $\langle u^*, MN^T u^* \rangle < 0$;

(iii) $u^* \neq 0$.

Proof. (i) Since (x^*, y^*) is an stationary point of (2.7), we can write

$$-2M^T u^* + \nabla_x \Phi(x^*, y^*) = 0 \tag{2.23}$$

$$2N^T u^* + \nabla_y \Phi(x^*, y^*) = 0. \tag{2.24}$$

Suppose on the contrary that $\Phi(x^*, y^*) = 0$, then it holds that

$$\phi(x_i, y_i) = 0 \quad \forall i \in [n].$$

If $(x_i^*, y_i^*, w_i) \neq 0$, then

$$\frac{\partial}{\partial x_i}\Phi(x^*, y^*) = 2\phi(x_i^*, y_i^*)\frac{\partial}{\partial x_i}\phi(x_i^*, y_i^*) = 0 \tag{2.25}$$

and

$$\frac{\partial}{\partial y_i}\Phi(x^*, y^*) = 2\phi(x_i^*, y_i^*)\frac{\partial}{\partial x_i}\phi(x_i^*, y_i^*) = 0. \tag{2.26}$$

Also, by Lemma 2.2.1, the above equations (2.25) and (2.26) hold when $(x_i^*, y_i^*, w_i) = 0$. Consequently, (2.23) and (2.24) reduce to

$$M^T u^* = 0 \quad \text{and} \quad N^T u^* = 0,$$

which are exactly the necessary and sufficient conditions for the global minimizer of the convex optimization problem

$$\min_{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n} \|\Pi_{\mathbb{R}^m_+}(-LMx + LNy + q)\|^2. \tag{2.27}$$

It turns out that (x^*, y^*) is a global solution of (2.7) with minimum value of zero, that is,

$$LMx^* - LNy^* \geq b.$$

Combining the above, we conclude that $(x^*, y^*) \in \mathrm{SOL}(w, M, N, \mathcal{P})$, contradicting the hypothesis.

(ii) Suppose, by contradiction, that

$$\langle u^*, MN^T u^* \rangle = \langle M^T u^*, N^T u^* \rangle \geq 0.$$

We deduce from (2.23)-(2.26) and part (iii) of Lemma 2.2.1 that

$$4\left\langle M^T u^*, N^T u^*\right\rangle = -\sum_{i=1}^{n} \frac{\partial}{\partial x_i}\Phi(x^*, y^*)\frac{\partial}{\partial y_i}\Phi(x^*, y^*) \leq 0.$$

Therefore,

$$\frac{\partial}{\partial x_i}\Phi(x^*, y^*)\frac{\partial}{\partial y_i}\Phi(x^*, y^*) = 0,\ \forall i \in [n].$$

It follows from part (iv) of Lemma (2.2.1) that

$$\Phi(x^*, y^*) = 0.$$

This leads to a contradiction with (i).

(iii) It follows immediately from (ii). □

Corollary 2.5.1. *Assume that* (x^*, y^*) *is a stationary point of* (2.7)*, and define* $u^* = L^T \Pi_{\mathbb{R}^m_+}(-Mx^* + Ny^* + q)$*. Then* $(x^*, y^*) \in SOL(w, M, N, \mathcal{P})$ *if any of the following statements hold:*

(i) $FEA(w, M, N, \mathcal{P}) \neq \emptyset$ *and* $\Phi(x^*, y^*) = 0$*;*

(ii) $FEA(w, M, N, \mathcal{P}) \neq \emptyset$ *and* $\left\langle u^*, MN^T u^*\right\rangle \geq 0$*;*

(iii) $u^* = 0$*.*

Now, it is easy to deduce the following

Corollary 2.5.2. *Assume that* (x^*, y^*) *is a stationary point of* (2.7) *and define* $u^* = L^T \Pi_{\mathbb{R}^m_+}(-Mx^* + Ny^* + q)$*. If* $\{M, N\}$ *is row sufficient, then* $(x^*, y^*) \in SOL(w, M, N, \mathcal{P})$*.*

Proof. From part (iii) of Lemma 2.2.1, we have

$$4(M^T u^*)_i (N^T u^*)_i = -\frac{\partial}{\partial x_i}\Phi(x^*, y^*)\frac{\partial}{\partial y_i}\Phi(x^*, y^*) \leq 0 \quad \forall i \in [n].$$

The row sufficiency of $\{M, N\}$ implies that

$$\frac{\partial}{\partial x_i}\Phi(x^*, y^*)\frac{\partial}{\partial y_i}\Phi(x^*, y^*) = 0 \quad \forall i \in [n].$$

In view of (iv) of Lemma (2.2.1), we conclude that

$$\Phi(x^*, y^*) = 0,$$

showing that $(x^*, y^*) \in \mathrm{SOL}(w, M, N, \mathcal{P})$ by part (i) of Corollary 2.5.1. □

The following result gives a sufficient condition for the existence of a stationary point of (2.7).

Corollary 2.5.3. *If the objective function F of* (2.7) *has a bounded level set and the pair $\{M, N\}$ is row sufficient, then the feasibility of* (2.2) *implies its solvability.*

Proof. Pick any $(\bar{x}, \bar{y}) \in FEA(w, M, N, \mathcal{P})$. By assumption, the level set

$$\mathcal{L}(\bar{x}, \bar{y}) := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : F(x, y) \leq F(\bar{x}, \bar{y})\}$$

is bounded, and hence compact by the continuity of F. By the Weirestrass theorem, we know that (2.7) admits an optimal solution which satisfies the KKT conditions. It follows from Corollary 2.5.2 that the KKT vector is a solution to (2.2). This completes the proof. □

2.6 The P_*-Property

A square matrix $M \in \mathbb{R}^{n\times n}$ is said to have the *sufficiency* if it is both column sufficient and row sufficient. Analogously, we say a matrix pair $\{M, N\}$ has the sufficiency property if it is both column sufficient and row sufficient with respect to a polyhedra $\mathcal{P}$.

A closely related notion is the P_*-matrix introduced in [42] by Kojima et al for the development of interior-point methods for solving LCPs. A matrix $M \in \mathbb{R}^{n\times n}$ is said to have the $P_*(\kappa)$-property if there exists $\kappa \geq 0$ such that for any $x \in \mathbb{R}^n$, it holds

$$(1 + 4\kappa) \sum_{i\in\mathcal{I}_+} x_i(Mx)_i + \sum_{i\in\mathcal{I}_-} x_i(Mx)_i \geq 0, \tag{2.28}$$

where

$$\mathcal{I}_+ = \{i \in [n] : x_i(Mx)_i > 0\}, \quad \mathcal{I}_- = \{i \in [n] : x_i(Mx)_i < 0\}.$$

A matrix is said to have the P_*-property if it has the $P_*(\kappa)$ property for some $\kappa \geq 0$. Thus, we can write

$$P_* = \cup_{\kappa\geq 0} P_*(\kappa).$$

It was showed in [42] that a P_*-matrix is column sufficient. Subsequently, Guu and Cottle [22] proved that a P_*-matrix is also row sufficient and therefore the class P_* is included in the class of sufficient matrices. Interestingly, later Väliaho [61] proved the reverse inclusion, thereby showing that P_* coincides with the class of sufficient matrices.

We generalize the P_*-property for non-square matrix pairs.

Definition 2.6.1. Given matrices $M, N \in \mathbb{R}^{m\times n}$ and a non-empty polyhedral $\mathcal{P} \subseteq \mathbb{R}^m$, we say that $\{M, N\}$ has the $P_*(\kappa)$-property with respect to $\mathcal{P}$ if the following condition holds:

$$Mx - Ny \in \mathcal{P} - \mathcal{P} \implies (1 + 4\kappa) \sum_{i\in\mathcal{I}_+} x_i y_i + \sum_{i\in\mathcal{I}_-} x_i y_i \geq 0, \tag{2.29}$$

where $\kappa \geq 0$ is a given constant called the handicap of the pair $\{M, N\}$, and

$$\mathcal{I}_+ = \{i \in [n] : x_i y_i > 0\}, \quad \mathcal{I}_- = \{i \in [n] : x_i y_i < 0\}.$$

The pair $\{M, N\}$ is said to have the P_*-property if it has the $P_*(\kappa)$ property for some $\kappa \geq 0$.

Clearly, the $P_*(0)$ property reduces to the monotonicity of $\{M, N\}$. When $m = n$ and $\mathcal{P}$ is singleton, we obtain the $P_*(\kappa)$-property studied in the context of the HLCP [30] . In that case, Väliaho's result still holds, i.e., $\{M, N\}$ has the P_*-property if and only if it has the column sufficiency and row sufficiency with respect to $\mathcal{P}$. The P_*-property always implies the column sufficiency of $\{M, N\}$. However, it is not clear whether it is related to the row sufficiency property in general.

The next result concerns the existence of solutions of the **wXLCP**$(w, M, N, \mathcal{P})$.

Proposition 2.6.1. *If the pair $\{M, N\}$ has the row sufficiency and the $P_*(\kappa)$ property with respect to $\mathcal{P}$, then the **wXLCP**$(w, M, N, \mathcal{P})$ is solvable if it is strictly feasible.*

Proof. The proof adapts the argument used in the proof of Theorem 2.4 in [44]. Let $(\bar{x}, \bar{y})$ be a strictly feasible point, and we denote by

$$f_w(x, y) = x^T y - \sum_{i=1}^{n} w_i \log x_i y_i \tag{2.30}$$

the objective function of (2.2). Consider the sublevel set

$$\mathcal{L}_w(\bar{x}, \bar{y}) = \{(x, y) \in FEA(w, M, N, \mathcal{P}) : f_w(x, y) \leq f_w(\bar{x}, \bar{y})\}. \tag{2.31}$$

First, we note that

$$(x, y) \in \mathcal{L}_w(\bar{x}, \bar{y}) \Longrightarrow x_i y_i > 0, \quad \forall i \in \overline{\mathcal{W}}.$$

Indeed, if $x_i y_i = 0$ for some $i \in \overline{\mathcal{W}}$, then $f_w(x, y) = \infty > f_w(\bar{x}, \bar{y})$, leading to a contradiction. It is easy to verify the inequality

$$\rho - \sigma \log \rho \geq \sigma - \sigma \log \sigma, \quad \forall \rho, \sigma > 0, \tag{2.32}$$

which implies that for any $(x, y) \in \mathcal{L}_w(\bar{x}, \bar{y})$ there holds

$$\begin{aligned} \sum_{i \in \widehat{\mathcal{W}}} x_i y_i &= f_w(x, y) - \sum_{i \in \overline{\mathcal{W}}} (x_i y_i - w_i \log x_i y_i) \\ &\leq f_w(\bar{x}, \bar{y}) - \sum_{i \in \overline{\mathcal{W}}} (w_i - w_i \log w_i). \end{aligned}$$

It follows that for any $j \in \overline{\mathcal{W}}$,

$$\begin{aligned} x_j y_j - w_j \log x_j y_j &= f_w(x, y) - \sum_{i \in \widehat{\mathcal{W}}} x_i y_i - \sum_{i \in \overline{\mathcal{W}} - \{j\}} (x_i y_i - w_i \log x_i y_i) \\ &\leq f_w(\bar{x}, \bar{y}) - \sum_{i \in \overline{\mathcal{W}} - \{j\}} (w_i - w_i \log w_i). \end{aligned}$$

Consequently, we deduce that there are constants λ and Λ satisfying

$$0 < \lambda \leq x_j y_j \leq \Lambda, \quad \forall j \in \overline{\mathcal{W}}. \tag{2.33}$$

Therefore, we obtain that $x^T y \leq \eta$ for some $\eta > 1$. Observe that

$$M(\bar{x} - x) - N(\bar{y} - y) \in \mathcal{P} - \mathcal{P},$$

we deduce from the $P_*(\kappa)$ property of $\{M, N\}$ that

$$\langle \bar{x} - x, \bar{y} - y \rangle \geq -4\kappa \sum_{i \in \mathcal{I}_+} (\bar{x} - x)_i (\bar{y} - y)_i,$$

where κ is the handicap, and

$$\mathcal{I}_+ = \{i \in [n] : (\bar{x} - x)_i (\bar{y} - y)_i > 0\}.$$

It follows that

$$\begin{aligned} \bar{x}^T y + \bar{y}^T x &\leq \bar{x}^T \bar{y} + x^T y + 4\kappa \sum_{i \in \mathcal{I}_+} (\bar{x}_i \bar{y}_i + x_i y_i - \bar{x}_i y_i - \bar{y}_i x_i) \\ &\leq 2\eta + 4\kappa \sum_{i \in \mathcal{I}_+} (\bar{x}_i \bar{y}_i + x_i y_i) \leq 2(1 + 4\kappa)\eta. \end{aligned}$$

This shows that each component of x and y is bounded for any $(x, y) \in \mathcal{L}_w(\bar{x}, \bar{y})$. That is, the sublevel set $\mathcal{L}_w(\bar{x}, \bar{y})$ is bounded. It is also closed by the continuity of f_w. We conclude that f_w has a minimum on the compact set $\mathcal{L}_w(\bar{x}, \bar{y})$. The solvability now follows from Proposition 2.5.1. This completes the proof. □

Corollary 2.6.1. *If the pair $\{M, N\}$ has the monotonicity with respect to the polyhedral $\mathcal{P}$, then the **wXLCP**$(w, M, N, \mathcal{P})$ is solvable whenever it is strictly feasible.*

2.7 The P-Property

We now consider the P-property. In the context of the LCP, for a matrix $M \in \mathbb{R}^{n \times n}$, the **LCP**$(M, q)$ has a unique solution for any $q \in \mathbb{R}^n$ if and only if M is a P-matrix, i.e.,

$$x \circ Mx \leq 0 \implies x = 0.$$

There are a number of different characterizations of the P-property, some of which will be summarized in Chapter 3.

In [17], Gowda extended the P-property for a matrix pair $\{M, N\}$ in the context of the XLCP.

Definition 2.7.1. Given matrices $M, N \in \mathbb{R}^{m\times n}$ and a non-empty polyhedral $\mathcal{P} \subseteq \mathbb{R}^m$, we say that $\{M, N\}$ has the P-property with respect to $\mathcal{P}$ if the following conditions hold:

$$\left.\begin{aligned} Mx - Ny \in \mathcal{P} - \mathcal{P} \\ x \circ y \leq 0 \end{aligned}\right\} \implies (x, y) = (0, 0). \tag{2.34}$$

$$\left.\begin{aligned} (M^T u) \circ (N^T u) \leq 0 \\ u \in (0^+\mathcal{P})^* \end{aligned}\right\} \implies u = 0. \tag{2.35}$$

When $m = n$ and $\mathcal{P}$ is a singleton, conditions (2.34) read

$$\left.\begin{aligned} Mx - Ny = 0 \\ x \circ y \leq 0 \end{aligned}\right\} \implies (x, y) = (0, 0). \tag{2.36}$$

$$(M^T u) \circ (N^T u) \leq 0 \implies u = 0. \tag{2.37}$$

It should be observed that (2.37) can be implied by (2.36). To see this, we first notice that (2.36) is satisfied if and only if M is nonsingular and $M^{-1}N$ is a P-matrix. Otherwise, there exists $\bar{x} \neq 0$ such that $M\bar{x} = 0$, then $(\bar{x}, 0)$ violates the implication (2.36). Now, consider any u satisfying $(M^T u) \circ (N^T u) \leq 0$. Denote $v = M^T u$, then $u = M^{-T} v$ and $v \circ N^T M^{-T} v \leq 0$, implying that $v = 0$ and hence $u = 0$ by the P-matricity of $M^{-1}N$.

The following result ties the P-property with the uniqueness of the XLCP.

Theorem 2.7.1 ([17])**.** *If the pair $\{M, N\}$ has the P-property with respect to $\mathcal{P}$ if and only if for every $p \in \mathbb{R}^m$ and every column rearrangement $\{\overline{M}, \overline{N}\}$ of $\{M, N\}$, the XLCP$(\overline{M}, \overline{N}, \mathcal{P} + p)$ has a unique solution.*

We aim to relate the P-property with the uniqueness of the wXLCP. First, we show that condition (2.35) implies the strict feasibility of the problem.

Lemma 2.7.1. *If condition* (2.35) *is satisfied, then the* ***wXLCP***$(w, M, N, \mathcal{P})$ *is strictly feasible.*

Proof. Assume that $\mathcal{P} = \{u \in \mathbb{R}^m : Lu \geq b\}$ for some $L \in \mathbb{R}^{k\times m}$ and $b \in \mathbb{R}^m$. It suffices to show that the following system has a solution

$$LMx - LNy > 0, \; x \geq 0, \; y \geq 0. \tag{2.38}$$

Indeed, if $(\bar{x}, \bar{y})$ satisfies (2.38) then, by continuity, the pair $(x, y) = (\bar{x} + \varepsilon e, \bar{y} + \varepsilon e)$ is a strictly feasible point, where $\varepsilon > 0$ is sufficiently small and e is the vector of all ones.

Suppose, on the contrary, that (2.38) is inconsistent so that the convex sets $A = \{u \in \mathbb{R}^m : Lu > 0\}$ and $B = M(\mathbb{R}^n_+) - N(\mathbb{R}^n_+)$ are disjoint. By a separation theorem [35], there exists a nonzero vector v and a scalar c such that

$$v^T u > c \geq v^T(Ms - Nt), \; \forall\, u \in A, \; s, t \in \mathbb{R}^n_+.$$

Note that $0 \in B$, we have $c \geq 0$, and thus $v^T u > c \geq 0$ for all $u \in A$, implying that $v \in (0^+\mathcal{P})^*$. Fix $t = 0$, we deduce from the above inequality that $v^T M s \leq c$ for all $s \in \mathbb{R}^n_+$. We claim that $M^T v \leq 0$. Otherwise, if $(M^T v)_j > 0$ for some j, by choosing $s = \lambda e_j$ with e_j being the j-th standard unit vector and $\lambda > 0$, then $v^T M s = \lambda (M^T v)_j \to \infty$ as $\lambda \to \infty$. The same reasoning implies that $N^T v \geq 0$. Therefore, the nonzero vector v satisfies the left-hand condition in (2.35), implying that $v = 0$. This leads to a contradiction. □

Proposition 2.7.1. *The following statements hold*

(i) If $\{M, N\}$ has the P and $P_*(\kappa)$-property with respect to $\mathcal{P}$, then ***wXLCP***$(w, M, N, \mathcal{P} + p)$ *has a unique solution for any $w \in \mathbb{R}^n_+$ and any $p \in \mathbb{R}^m$.*

(ii) If ***wXLCP***$(w, \overline{M}, \overline{N}, \mathcal{P} + p)$ *has a unique solution for every column rearrangement $\{\overline{M}, \overline{N}\}$ of $\{M, N\}$, for every $w \in \mathbb{R}^n_+$ and every $p \in \mathbb{R}^m$, then $\{M, N\}$ has the P-property with respect to $\mathcal{P}$.*

Proof. (i). The existence of solutions follows from Proposition 2.6.1 and Lemma 2.7.1. For the uniqueness, we assume that $(\bar{x}, \bar{y})$ and $(\hat{x}, \hat{y})$ are two solutions. Let $(u, v) = (\bar{x} - \hat{x}, \bar{y} - \hat{y})$, easy verification shows that (u, v) satisfies the left-hand side condition in (2.34), thus $(u, v) = (0, 0)$. This shows that $(\bar{x}, \bar{y}) = (\hat{x}, \hat{y})$ as desired.

(ii). This directly follows from Theorem (2.7.1). □

2.8 P_0 and R_0 Properties

First, let's recall the P_0 and R_0 properties in the context of the LCP.

Definition 2.8.1. The matrix $M \in \mathbb{R}^{n \times n}$ is said to be

- a P_0-matrix if for any $0 \neq x \in \mathbb{R}^n$, there is an index i with $x_i \neq 0$ such that $x_i (Mx)_i \geq 0$.
- a R_0-matrix if
$$x \geq 0,\ Mx \geq 0,\ \langle x, Mx \rangle = 0 \implies x = 0.$$

Now, we formulate these notions for matrix pairs.

Definition 2.8.2. Given matrices $M, N \in \mathbb{R}^{m \times n}$ and a polyhedral $\mathcal{P} \subseteq \mathbb{R}^m$. We say that the pair $\{M, N\}$ has the

- P_0-property with respect to $\mathcal{P}$ if the following implication holds:
$$\left.\begin{aligned} Mx - Ny \in \mathcal{P} - \mathcal{P} \\ (x, y) \neq 0 \end{aligned}\right\} \implies x_i y_i \geq 0 \text{ for some } i \text{ such that } |x_i| + |y_i| > 0. \tag{2.39}$$

- R_0-property with respect to $\mathcal{P}$ if the implication holds:

$$\left.\begin{aligned} Mx - Ny \in \mathcal{P} - \mathcal{P} \\ (x, y) \geq 0, \ \ \langle x, y \rangle = 0 \end{aligned}\right\} \implies (x, y) = (0, 0). \tag{2.40}$$

We relate the definitions above to existing concepts.

(i) When $m = n$ and $\mathcal{P} = \{q\}$, condition (2.39) implies

$$\left.\begin{aligned} Mx - Ny = 0 \\ (x, y) \neq 0 \end{aligned}\right\} \implies x_i y_i \geq 0 \quad \text{for some } i \text{ such that } |x_i| + |y_i| > 0. \tag{2.41}$$

In addition, if N is the identity matrix, (2.41) reduces to the P_0-property of M.

Though a row sufficient matrix is also a P_0-matrix, we show in the following example that this relationship does not hold between (2.10) and (2.39). Take

$$M = N = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Straightforward calculations show that $\{M, N\}$ has the row sufficiency property with respect to singleton. To show that the P_0-property of $\{M, N\}$ is violated, we consider $x = (0, 1)^T$ and $y = (0, -1)^T$, which clearly satisfy the left-hand side condition in (2.39). However, there is no index $i \in \{1, 2\}$ such that $|x_i| + |y_i| > 0$ and $x_i y_i \geq 0$.

(ii) When $m = n$ and $\mathcal{P} = \{q\}$, condition (2.40) simplifies as

$$\left.\begin{aligned} Mx - Ny = 0 \\ (x, y) \geq 0, \ \ \langle x, y \rangle = 0 \end{aligned}\right\} \implies (x, y) = (0, 0). \tag{2.42}$$

It can be easily shown that (2.42) holds if and only if the associated HCLP has a bounded solution set. In particular, if N is the identity matrix, (2.42) is equivalent to saying that M is a R_0-matrix. We show that condition (2.40) is also related to the boundedness of certain wXLCPs.

Lemma 2.8.1. *Let $M, N \in \mathbb{R}^{m \times n}$ and $\mathcal{P} \subseteq \mathbb{R}^m$ be a non-empty polyhedral. If $\{M, N\}$ has the R_0-property with respect to $\mathcal{P}$, then $SOL(w, M, N, \mathcal{P})$ is a bounded (possibly empty) if any of the following conditions is satisfied:*

(a) $\mathcal{P}$ is bounded.

(b) $\mathcal{P}$ is a closed convex cone.

(c) $\mathcal{P}$ is an affine subspace.

Proof. Suppose that for some $w \in \mathbb{R}^n_+, p \in \mathbb{R}^m$ there is an unbounded sequence $(x^k, y^k)_{k=1}^{\infty} \subset \mathrm{SOL}(w, M, N, \mathcal{P})$. Then, for each k, there exists $b^k \in \mathcal{P}$ such that $\|(x^k, y^k)\| \to \infty$ and

$$Mx^k - Ny^k = b^k + p, \ x^k \circ y^k = w, \ (x^k, y^k) \geq 0. \tag{2.43}$$

(a). If $\mathcal{P}$ is bounded, we have that $\{b^k\}$ is bounded. Dividing both sides of (2.43) by $\|(x^k, y^k)\|$, and by taking a subsequence and passing to the limit if necessary, we deduce that there exists $(\bar{x}, \bar{y})$ with $\|(\bar{x}, \bar{y})\| = 1$ such that

$$M\bar{x} - N\bar{y} = 0 \in \mathcal{P} - \mathcal{P}, \ \bar{x} \circ \bar{y} = 0, \ (\bar{x}, \bar{y}) \geq 0,$$

implying that $(\bar{x}, \bar{y}) = 0$ by the R_0 property of $\{M, N\}$. This leads to a contradiction.

(b). If $\mathcal{P}$ is a closed convex cone, then $0 \in \mathcal{P}$ and $\mathcal{P}$ is invariant under a positive scaling. Thus the normalization procedure yields

$$M\bar{x} - N\bar{y} \in \mathcal{P} \subset \mathcal{P} - \mathcal{P}, \ \bar{x} \circ \bar{y} = 0, \ (\bar{x}, \bar{y}) \geq 0,$$

which implies that $(\bar{x}, \bar{y}) = 0$, leading to a contradiction.

(3). Let $\mathcal{P}$ be a subspace, i.e., $\mathcal{P} = \{u \in \mathbb{R}^m : Lu = q\}$, where $q \in \mathbb{R}^m$ and $L \in \mathbb{R}^{k \times m}$ has a full row rank. Then $\mathcal{P} - \mathcal{P} = \{u \in \mathbb{R}^m : Lu = 0\}$ and (2.43) gives

$$LMx^k - LNy^k = q + Lp, \ x^k \circ y^k = w, \ (x^k, y^k) \geq 0. \tag{2.44}$$

and the R_0-property of $\{M, N\}$ simplifies to

$$\left.\begin{array}{r} LMx - LNy = 0 \\ (x, y) \geq 0, \ \langle x, y \rangle = 0 \end{array}\right\} \implies (x, y) = (0, 0). \tag{2.45}$$

It is worth noting that (2.45) implies

$$\left.\begin{array}{r} Mx - Ny = 0 \\ (x, y) \geq 0, \ \langle x, y \rangle = 0 \end{array}\right\} \implies (x, y) = (0, 0), \tag{2.46}$$

Again, the resulting normalized system obtained from (2.44) will contradict condition (2.46). This completes the proof. □

Recall that the LCP(q, M) has an nonempty solution set if M is both P_0 and R_0. We extend this result to certain wXLCPs. To this end, we consider the mapping $G : \mathbb{R}^n_+ \times \mathbb{R}^n_+ \to \mathbb{R}^n_+ \times \mathbb{R}^n$ defined by

$$G(x, y) = \begin{pmatrix} x \circ y \\ H(x, y) \end{pmatrix}, \tag{2.47}$$

where $H : (x, y) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+ \mapsto \mathbb{R}^n$. The map (2.47) is a special case of the the interior mapping studied in [39], which plays a fundamental role in the family of path-following interior-point methods for solving complementarity problems. Its properties have been examined in the contexts of NCP, the HLCP and the SDCP [39]. According to Theorem 2 and 3 in [39], when H is specialized to the linear mapping $Mx - Ny$, G has the following properties.

Proposition 2.8.1. *Let $M, N \in \mathbb{R}^{n\times n}$ and $H(x,y) = Mx - Ny$. If conditions* (2.41) *and* (2.42) *are satisfied, then the following statements hold*

(1) G maps $\mathbb{R}^n_+ \times \mathbb{R}^n_+$ homeomorphically onto $\mathbb{R}^n_+ \times \mathbb{R}^n$;

(2) $G(\mathbb{R}^n_+ \times \mathbb{R}^n_+) = \mathbb{R}^n_+ \times \mathbb{R}^n$.

Combing Proposition 2.8.1 and Lemma 2.8.1, we have the following existence result of the wXLCP.

Proposition 2.8.2. *Let $M, N \in \mathbb{R}^{n\times n}$ and $\mathcal{P}$ be a non-empty polyhedral in $\mathbb{R}^n$. If $\{M, N\}$ has the P_0 and R_0 properties with respect to $\mathcal{P}$, then the **wXLCP**$(w, M, N, \mathcal{P})$ has a solution. Moreover, the set SOL$(w, M, N, \mathcal{P})$ is bounded if any of the conditions (a)-(c) in Lemma 2.8.1 is satisfied.*

As a consequence of the above result, we may apply a standard homotopy method to find a solution of **wHLCP**(w, M, N, q), i.e.,

$$\begin{cases} x \circ y = w \\ Mx - Ny = q \\ (x,y) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+. \end{cases}$$

Consider the system of equations parametrized by the parameter $t \in [0, 1]$:

$$\begin{cases} x \circ y = (1-t)w + ta \\ Mx - Ny - q = tp \\ (x,y) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+. \end{cases} \tag{2.48}$$

where $a = x^0 \circ y^0$ and $p = Mx^0 - Ny^0 - q$ for a given $(x^0, y^0) \in \mathbb{R}^n_{++} \times \mathbb{R}^n_{++}$. It is easily seen that when $t = 1$, the paramerized system has the solution (x^0, y^0). On the other hand, when $t = 0$, it reduces to our original problem of interest. We have from Proposition 2.8.1, for each $t \in (0, 1]$, (2.48) has a unique solution $(x(t), y(t)) \in \mathbb{R}^n_{++} \times \mathbb{R}^n_{++}$. Moreover, assuming the conditions in Lemma 2.8.1, the set $\{(x(t), y(t)) : t \in (0, 1]\}$ is bounded hence forms a continuous path, and every limit point of the path, as $t \downarrow 0$, is a solution of **wHCLP**(w, M, N, q).

Corollary 2.8.1. *Let $\mathcal{P}$ be an affine space. If the pair $\{M, N\}$ has the R_0 and row sufficiency properties with respect to $\mathcal{P}$, then the feasibility of* (2.2) *implies its solvability.*

Proof. By Corollary 2.5.3, it suffices to show that F has the bounded level set property. Suppose, on the contrary, that there exists an unbounded sequence (x^k, y^k) such that

$$F(x^k, y^k) \le \alpha$$

for some $\alpha > 0$. Applying the standard normalization technique, we obtain $(\bar{x}, \bar{y}) \geq 0$ with $\|(\bar{x}, \bar{y})\| = 1$ satisfying

$$M\bar{x} - N\bar{y} = 0, \ \bar{x} \circ \bar{y} = 0.$$

This contradicts the R_0-property of $\{M, N\}$. □

We summarize the connections between properties in the following result.

Proposition 2.8.3. *Let $M, N \in \mathbb{R}^{m \times n}$ and $\mathcal{P}$ be non-empty a polyhedral in $\mathbb{R}^m$. The following statements hold:*

(i) Monotonicity property $\Longrightarrow$ sufficiency property.

(ii) P_*-property $\Longrightarrow$ Column sufficiency $\Longrightarrow$ P_0-property.

(iii) P-property $\Longrightarrow$ sufficiency property.

(iv) P-property $\Longrightarrow$ R_0-property.

Proof. (i) The column sufficiency follows immediately from the definitions, and the row sufficiency is a consequence of Proposition 2.3.1 and 2.5.1.

(ii) Only the second inclusion requires a proof. Suppose that the pair $\{M, N\}$ does not have the P_0-property, then there exists a positive diagonal matrix D and a pair $(\bar{x}, \bar{y})$ with $|\bar{x}_j| + |\bar{y}_j| > 0$ such that

$$M\bar{x} - N\bar{y} \in \mathcal{P} - \mathcal{P}, \quad D\bar{x} + \bar{y} = 0.$$

Note that $\bar{x} \circ \bar{y} = \bar{x} \circ (-D\bar{x}) \leq 0$, so $\bar{x} \circ \bar{y} = 0$ by the column sufficiency of $\{M, N\}$. However, we have $\bar{x}_j \bar{y}_j = -D_{jj} \bar{x}_j^2 < 0$, leading to a contradiction.

The claims (iii) and (iv) are trivial. □

2.9 Extension

Similar to standard complementarity problems, the weighted complementarity problem can also be defined over a Euclidean Jordan algebra. Let $(\mathbb{E}, \langle \cdot, \cdot \rangle, \circ)$ be a finite dimensional Euclidean Jordan algebra over $\mathbb{R}$ and $\mathcal{K} \subset \mathbb{E}$ be a symmetric cone, the weighted mixed complementarity problem (wMCP) is a problem of finding $(x, y, z) \in \mathcal{K} \times \mathcal{K} \times \mathbb{R}^m$ such that

$$F(x, y, z) = 0, \quad x \circ y = w, \tag{2.49}$$

where $F : \mathbb{E} \times \mathbb{E} \times \mathbb{R}^m \to \mathbb{E} \times \mathbb{R}^m$ is a continuous nonlinear transformation, and $w \in \mathcal{K}$ is a given vector. When the function F is affine, i.e., $F(x, y, z) = Ax + By + Cz - a$, where a is a vector in $\mathbb{E} \times \mathbb{R}^m$, and $A : \mathbb{E} \to \mathbb{E} \times \mathbb{R}^m, B : \mathbb{E} \to \mathbb{E} \times \mathbb{R}^m, C : \mathbb{E} \to \mathbb{E} \times \mathbb{R}^m$ are linear transformations, we call the problem (2.49) the weighted mixed linear complementarity problem. When the weight vector $w = 0$, then the *wMCP* reduces to the *mixed complementarity problem over symmetric cones*, or the *mixed SCCP*, which contains the various standard symmetric cone complementarity problems (SCCP) discussed in last chapter as special cases.

Our main goal of studying (2.49) is to study its solution properties. An existence result was implicitly established in [66], where a number of properties of the interior point mapping

$$H : (x, y, z) \in \mathcal{K} \times \mathcal{K} \times \mathbb{R}^m \mapsto \begin{pmatrix} x \circ y \\ F(x, y, z) \end{pmatrix}$$

were derived for analyzing the solution properties of the mixed SCCP. More specifically, under the following technical conditions on the mapping F:

(A1) F is (x, y)-equilevel-monotone, i.e., for any $(x, y, z), (x', y', z') \in \mathbb{E} \times \mathbb{E} \times \mathbb{R}^m$

$$F(x, y, z) = F(x', y', z') \Longrightarrow \langle x - x', y - y' \rangle \geq 0,$$

(A2) F is z-injective, i.e., for any $(x, y, z), (x', y', z') \in \mathbb{E} \times \mathbb{E} \times \mathbb{R}^m$,

$$F(x, y, z) = F(x', y', z') \Longrightarrow z = z',$$

(A3) F is z-bounded, i.e., for every sequence $\{(x^k, y^k, z^k)\} \subset \mathbb{E} \times \mathbb{E} \times \mathbb{R}^m$ such that $\{(x^k, y^k)\}$ and $\{F(x^k, y^k, z^k)\}$ are bounded, the sequence $\{z^k\}$ is also bounded,

then it holds that

$$H(\mathcal{K} \times \mathcal{K} \times \mathbb{R}^m) \supset \mathcal{K} \times F(\mathrm{int}(\mathcal{K}) \times \mathrm{int}(\mathcal{K}) \times \mathbb{R}^m).$$

Therefore, if F satisfies the aforementioned assumptions (A1)-(A3) and the wMCP (2.49) is strictly feasible, then it is solvable for any $w \in \mathcal{K}$.

In future work, we hope to derive more results of the wMCP from both the theoretical and algorithmic point of view.

Chapter 3

Uniform Non-singularity Property

3.1 Motivation

There recently has been much work on numerical methods for solving the **SCCP**$(F, \mathcal{K})$ (cf. (1.8)), including smoothing Newton methods [2, 46, 56, 32, 64, 54, 55], interior-point method [49], and non-interior continuation methods [4, 25], to name a few. It is noted that a common feature among these references is the assumption of the monotonicity on F when $\mathcal{K}$ is different from $\mathbb{R}^n_+$. This assumption plays a critical role in algorithm designs since it provides a sufficient condition to guarantee the existence of Newton direction as well as the boundedness of the iterates encountered in those numerical approaches. A question which is of general theoretical interest is whether one can extend the existing well-developed algorithms for more general problems other than monotone problems.

In the theory of the LCP, a class of nonmonotone problems which has been well documented is the P-LCPs, i.e., the matrix M is a P-matrix [43]. Recall that a matrix $M \in \mathbb{R}^{n\times n}$ is a P-matrix if any of the following equivalent characterizations holds:

(1) All of its principal minors are positive [14].

(2) The following implication holds [16]

$$x \in \mathbb{R}^n, x \circ (Mx) \leq 0 \implies x = 0.$$

(3) The **LCP**(M, q) is globally uniquely solvable for all $q \in \mathbb{R}^n$ [60].

(4) The function $F(x) := M\Pi_{\mathbb{R}^n_+}(x) + x - \Pi_{\mathbb{R}^n_+}(x)$ is invertible in a neighborhood of zero [60].

(5) The function $F(x) := M\Pi_{\mathbb{R}^n_+}(x) + x - \Pi_{\mathbb{R}^n_+}(x)$ is invertible in a neighborhood of zero with Lipschitzian inverse.

(6) The matrix $DM + I - D$ is non-singular for any diagonal matrix D with $0 \leq D \leq I$, i.e., $0 \leq D_{ii} \leq 1$ for all $i \in [n]$. [33].

(7) The matrix $DM + I - D$ is a P-matrix for any diagonal matrix D with $0 \leq D \leq I$ [21] .

Note that (3) gives a characterizing condition for the P-matrix in terms of the solution property of the associated LCP. The equivalence of (4) and (5) follows from the fact that the inverse of a piecewise affine function is also piecewise affine and hence Lipschitzian. Based on the characterization (6), Chen and Xiang [6, 5] derived an sharper error bound for P-LCP. From (6) and (7), we can conclude that M is a P-matrix if and only if $\det(DM + I - D) > 0$ for all diagonal D with $0 \leq D \leq I$.

Numerical approaches for nonmonotone NCPs have also been extensively studied based on the concept of P-function. See, e.g., [11]. Accordingly, there has been some considerable effort to extend the concept of P-property to transformations over a general Euclidean space. Generalized from property (2), a so called P-type property was introduced by Gowda and Song [19] in identifying a class of nonmonotone semidefinite linear complementarity problems (SDLCPs). Later on, they extended this property for transformations over Euclidean Jordan algebras [57, 59]. It was shown there that the analog of (3) $\Rightarrow$ (2) holds in that setting. It is, however, not clear whether any numerical algorithms can be designed based on their P-type properties. In the paper [3], Chen and Qi proposed the concept of Cartesian P-property and showed that the merit function approach and smoothing method can be applied to a class of nonmonotone SDLCPs, namely Cartesian P-SDLCPs. The natural extension of the Cartesian P-property to the case where $\mathcal{K}$ is a symmetric cone is

$$\max_{1 \leq \nu \leq \kappa} \langle x_\nu, L(x)_\nu \rangle > 0 \quad \forall x \neq 0, \tag{3.1}$$

where L is a linear transformation on $\mathbb{E}$, x_ν denotes the ν-th component of x in the direct sum $\mathbb{E} = \mathbb{E}_1 \oplus \cdots \oplus \mathbb{E}_\kappa$ of Euclidean Jordan algebras corresponding to the direct sum $\mathcal{K} = \mathcal{K}_1 \oplus \cdots \oplus \mathcal{K}_\kappa$ of irreducible symmetric cones. In one extreme case where the $\mathbb{E}_\nu$'s are isomorphic to $\mathbb{R}$, the Cartesian P-property reduces to the P-property of the matrix representation of L (i.e., property (2)). Though several impressive results such as the globally unique solvability property and the local Lipschitzian property of the solution map can be carried over from LCPs to Cartesian P-SDLCPs, the Cartesian P-property differs from the strong monotonicity only when $\mathcal{K}$ is a symmetric cone that can be written as a direct sum $\mathcal{K}_1 \oplus \cdots \oplus \mathcal{K}_\kappa$ of at least two irreducible symmetric cones.

In a recent paper by Chua and Yi [8], a new characterization of the P-matrix as being uniformly nonsingular under the addition of nonnegative diagonal D was used as the basis of definition of a P-type property for transformations on Euclidean Jordan algebras. Based on the P-type property, a continuation method with global convergence was studied for the $\mathbf{SCCP}(F, \mathcal{K})$. It was shown that the P-type property is weaker than the Cartesian P-property in general; and it lies between the concept of P-function and uniform P-function when $\mathcal{K} = \mathbb{R}^n_+$. Therefore, the uniform nonsingularity property gives a wider class of nonmonotone SCCPs that can be solved numerically.

In this chapter, we introduce the notion of *uniform non-singularity property* over some set $\mathcal{L}$ of adjoint linear transformations. We show that, as $\mathcal{L}$ varies, the property is closely related to several aforementioned concepts, including the P-matrix, strong monotonicity, a weaker version of the Cartesian P-property and the P-type property in [8].

3.2 Uniform Nonsingularity

We first formulate the concept of uniform nonsingularity for transformations over a finite dimensional Euclidean space $\mathbb{E}$.

Definition 3.2.1. A transformation $F : \mathbb{E} \to \mathbb{E}$ is said to be *uniformly nonsingular with modulus* α over a set $\mathcal{L}$ of self-adjoint linear transformations if $\alpha > 0$ and

$$\forall x, y \in \mathbb{E},\ \forall D \in \mathcal{L},\ \|F(x) - F(y) + D(x - y)\| \geq \alpha\|x - y\|. \tag{3.2}$$

Functions over $\mathbb{R}^n$

To illustrate Definition 3.2.1, we begin with the case where $\mathbb{E} = \mathbb{R}^n$ and $\mathcal{L}$ is the set of nonnegative diagonal matrices. Then the condition in (3.2) can be described as the following property.

Property 3.2.1. There is a constant $\alpha > 0$ such that for any nonnegative diagonal matrix $D \in \mathbb{R}^{n\times n}$ and any $x, y \in \mathbb{R}^n$,

$$\|F(x) - F(y) + D(x - y)\| \geq \alpha\|x - y\|. \tag{3.3}$$

Note that if $F(x) = Mx + q$ for some $M \in \mathbb{R}^{n\times n}$ and $q \in \mathbb{R}^n$, Property 3.2.1 further reduces to

$$\|Mx + Dx\| \geq \alpha\|x\| \tag{3.4}$$

for all $x \in \mathbb{R}^n$ and all nonnegative diagonal matrix D. It was shown in [8] that the implication (3.4), being a special case of a much broader result, yields a new characterization of the P-matrix in terms of the norm. The proof follows from a perturbation argument.

The Property 3.2.1 is closely related to the concept of P-functions which have been extensively studied in the context of NCPs.

Definition 3.2.2. The function $F : \mathbb{R}^n \to \mathbb{R}^n$ is said to be

- a P-function if for all $x, y \in \mathbb{R}^n$, $x \neq y$, there is an index $i \in [n]$ such that

$$(x_i - y_i)(F_i(x) - F_i(y)) > 0;$$

- a uniform P-function with modulus $\gamma > 0$ if for all $x, y \in \mathbb{R}^n$ there is an index $i \in [n]$ such that

$$(x_i - y_i)(F_i(x) - F_i(y)) \geq \gamma\|x - y\|^2.$$

The following inclusions have been established in [8].

Proposition 3.2.1. *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous function. It holds that*

(a) If F is a uniform P-function, then it satisfies the Property 3.2.1.

(b) If F satisfies the Property 3.2.1, then it is a P-function.

In general, the P-property of a nonlinear function F does not ensure the existence of solutions to the NCP$(F, \mathbb{R}^n_+)$.

Example 3.2.1. *Let $F : \mathbb{R}^2 \to \mathbb{R}^2$ be defined as*

$$F(x_1, x_2) = (-e^{-x_1} + x_2, x_2).$$

Since the Jacobian matrix of F at any $x \in \mathbb{R}^2$ is

$$\begin{bmatrix} e^{-x_1} & 0 \\ 1 & 1 \end{bmatrix}$$

which is a P-matrix, it follows that F is a P-function. Also, it is easy to see taht the feasible set of NCP$(F, \mathbb{R}^n_+)$ is nonempty. However, suppose we have for some $(x_1, x_2) \geq 0$ satisfying

$$x_1(-e^{-x_1} + x_2) = x_2^2 = 0 \Longrightarrow (x_1, x_2) = (0, 0),$$

which is not a feasible point. Thus, the NCP$(F, \mathbb{R}^n_+)$ has no solution.

It is well known that NCP$(F, \mathbb{R}^n_+)$ has a unique solution if F is a continuous uniform P-function [23, Theorem 3.9]. The conclusion still holds when F satisfies Property 3.2.1.

Proposition 3.2.2 ([8])**.** *If F is a continuous function satisfying Property 3.2.1, then the NCP$(F, \mathbb{R}^n_+)$ has a unique solution.*

From above results, it seems that functions satisfying Property 3.2.1 behave like uniform P-functions. Therefore, a natural question is: Are uniform P-property functions characterized by Property 3.2.1? A counterexample given in [8] shows that the answer is negative. Thus, Property 3.2.1 lies strictly between the P-function and the uniform P-function.

Next, we show that a strengthening of Property 3.2.1 gives a new characterization of the strong monotonicity.

Property 3.2.2. There is a constant $\alpha > 0$ such that for any positive semidefinite matrix $D \in \mathbb{S}^n_+$ and for any $x, y \in \mathbb{R}^n$

$$\|F(x) - F(y) + D(x - y)\| \geq \alpha \|x - y\|. \tag{3.5}$$

Definition 3.2.3. A transformation $F : \mathbb{E} \to \mathbb{E}$ is said to be *strongly monotone with modulus* $\gamma > 0$ if for any $x, y \in \mathbb{E}$

$$\langle F(x) - F(y), x - y\rangle \geq \gamma \|x - y\|^2.$$

When F satisfies the above inequality with $\gamma = 0$, we say that F is *monotone*.

Proposition 3.2.3. *The nonlinear function $F : \mathbb{R}^n \to \mathbb{R}^n$ is strongly monotone if and only if it satisfies Property 3.2.2.*

Proof. *"Only if":* Suppose that F is strongly monotone with modulus γ. For any $x, y \in \mathbb{R}^n$ and any positive semidefinite matrix $D \in \mathbb{S}^n_+$, it follows

$$\begin{aligned}\langle F(x) - F(y) + D(x - y), x - y\rangle &= \langle F(x) - F(y), x - y\rangle + \langle D(x - y), x - y\rangle \\ &\geq \langle F(x) - F(y), x - y\rangle \\ &\geq \gamma\|x - y\|^2.\end{aligned}$$

On the other hand, by the Cauchy-Schwarz's inequality, we have

$$\langle F(x) - F(y) + D(x - y), x - y\rangle \leq \|F(x) - F(y) + D(x - y)\|\|x - y\|.$$

Thus F satisfies Property 3.2.2.

"If": Now assume that F satisfies Property 3.2.2. Pick arbitrary $x, y \in \mathbb{R}^n$ with $x \neq y$. Let

$$z = \frac{\langle F(x) - F(y), x - y\rangle}{\|x - y\|^2}(x - y) - (F(x) - F(y)),$$

which satisfies $\langle z, x - y\rangle = 0$, and define, for each positive integer k, self-adjoint linear transformation

$$D_k : w \mapsto \frac{\langle w, kz + x - y\rangle}{k\|x - y\|^2}(kz + x - y),$$

which is monotone and satisfies $D_k(x-y) = z + \frac{x-y}{k}$. Under the assumption on F, we deduce

$$\alpha\|x - y\| \leq \|F(x) - F(y) + D_k(x - y)\| = \left\|\frac{\langle F(x) - F(y), x - y\rangle}{\|x - y\|^2}(x - y) + \frac{x - y}{k}\right\|.$$

Letting $k \to \infty$, the above inequality reduces to

$$|\langle F(x) - F(y), x - y\rangle| \geq \alpha\|x - y\|^2 > 0. \tag{3.6}$$

Since the self-adjoint linear transformation

$$D : w \mapsto -\frac{\langle w, F(x) - F(y)\rangle}{\langle F(x) - F(y), x - y\rangle}(F(x) - F(y))$$

satisfies

$$F(x) - F(y) + D(x - y) = 0,$$

Property 3.2.2 implies that $D \notin \mathbb{S}^n_+$, that is, $\langle F(x) - F(y), x - y\rangle > 0$. Hence, (3.6) reads

$$\langle F(x) - F(y), x - y\rangle \geq \alpha\|x - y\|^2.$$

Since x and y are arbitrary, we conclude that F is strongly monotone with modulus α. □

It should be pointed out that If we replace $\mathbb{R}^n$ and $\mathbb{S}^n_+$ with $\mathbb{E}$ and the set of self-adjoint monotone linear transformations over $\mathbb{E}$, respectively, the above result still follows from exactly the same argument.

Corollary 3.2.1. *Let $F : \mathbb{E} \to \mathbb{E}$ be a nonlinear transformation, then F is monotone if for any self-adjoint and monotone linear transformation D over $\mathbb{E}$, it holds that*

$$F(x) - F(y) + D(x - y) = 0 \Longrightarrow x = y.$$

Proof. It follows from the argument in the "if " part of the proof for Proposition 3.2.3. □

Lemma 3.2.1. *Let $M \in \mathbb{R}^{n\times n}$. The following statements are equivalent:*

(a) There exists $\alpha > 0$ such that

$$\|(MD + I - D)x\| \geq \alpha \|x\|$$

for all $x \in \mathbb{R}^n$ and all $D \in \mathbb{S}^n_+$ with $\|D\| \leq 1$.

(b) There exists $\beta > 0$ such that

$$\left\|(M + D^{-1} - I)x\right\| \geq \beta \|x\|$$

for all x and all $D \in \mathbb{S}^n_{++}$ (positive definite matrices) with $\|D\| \leq 1$.

(c) There exists $\gamma > 0$ such that

$$\|(DM + I - D)x\| \geq \gamma \|x\|$$

for all $x \in \mathbb{R}^n$ and all $D \in \mathbb{S}^n_+$ with $\|D\| \leq 1$.

Proof. (a) $\Longrightarrow$ (b): Let $D \in \mathbb{S}^n_{++}$ with $\|D\| \leq 1$. For any $x \in \mathbb{R}^n$, there exists $y \in \mathbb{R}^n$ such that $x = Dy$. Then

$$\left\|(M + D^{-1} - I)x\right\| = \|(MD + I - D)y\| \geq \alpha\|y\| = \alpha\|D^{-1}x\| \geq \alpha\|x\|,$$

where the last inequality used the fact that $\|D\| \leq 1$.

(b) $\Longrightarrow$ (a): Firstly, we assume that $D \in \mathbb{S}^n_{++}$. Fix $x \in \mathbb{R}^n$ and let $y = Dx$. We consider the following two cases.

If $\|y\| \geq \frac{\|x\|}{\beta+1+\|M\|}$, we have

$$\|(MD + I - D)x\| = \|(M + D^{-1} - I)y\| \geq \beta\|y\| \geq \frac{\beta}{\beta + 1 + \|M\|}\|x\|.$$

If $\|y\| < \frac{\|x\|}{\beta+1+\|M\|}$, then

$$\|(MD + I - D)x\| \geq \|x\| - \|y\| - \|My\| \geq \frac{\beta}{\beta + 1 + \|M\|}\|x\|.$$

Suppose now that $M \in \mathbb{S}^n_+$ is singular. Consider the nonsingular sequence $D^k = (1-\frac{1}{k})D+\frac{1}{k}I$. Since the cone $\mathbb{S}^n_+$ is convex, it follows that $D^k \in \mathbb{S}^n_{++}$. We then conclude from the above argument that

$$\|(MD + I - D)x\| = \lim_{k\to\infty} \|(MD^k + I - D^k)x\| \geq \frac{\beta}{\beta + 1 + \|M\|}\|x\|.$$

(c) $\Longleftrightarrow$ (a): Observe that (c) is the same as

$$\|(M^T D + I - D)x\| \geq \gamma\|x\|$$

for any $D \in \mathbb{S}^n_+$ with $\|D\| \leq 1$, which is equivalent to

$$\|(M^T + D^{-1} - I)x\| \geq \beta\|x\|$$

for all $D \in \mathbb{S}^n_{++}$ with $\|D\| \leq 1$ by the equivalence between (a) and (b). Taking the transpose once again, we get

$$\|(M + D^{-1} - I)x\| \geq \beta\|x\|,$$

which is in turn equivalent to (a). □

Corollary 3.2.2. *Let $M \in \mathbb{R}^{n\times n}$. The following statements are equivalent:*

(a) M is strongly monotone (not necessarily symmetric), i.e., $x^T Mx > 0$ for all $0 \neq x \in \mathbb{R}^{n\times n}$.

(b) There exists $\alpha > 0$ such that

$$\|(M + D)x\| \geq \alpha\|x\|,$$

for any $D \in \mathbb{S}^n_+$ and any $x \in \mathbb{R}^n$.

(c) The matrix $DM + I - D$ is non-singular for any $D \in \mathbb{S}^n_+$ with $\|D\| \leq 1$.

Proof. The equivalence between (a) and (b) follows from Proposition 3.2.3.

(b) $\iff$ (c): First, we note that

$$\mathbb{S}^n_+ = \{D^{-1} - I : \ D \in \mathbb{S}^n_{++} \text{ with } \|D\| \leq 1\}.$$

Also, using the fact that the set $\{D \in \mathbb{S}^n_+ : \|D\| \leq 1\}$ is compact, the statement (c) is equivalent to the statement (c) in 3.2.1. Now the equivalence between (b) and (c) follows from the equivalence between (b) and (c) in Lemma 3.2.1. □

Note that we can regard the Corollary 3.2.2 as the analog of (2) $\iff$ (6) for the P-matrix characterization in the introductory section. However, even though M is positive definite, the analog of (2) $\iff$ (7) does not hold. This can be illustrated by the following simple counterexample. Let

$$M = \begin{pmatrix} 10 & 3 \\ 3 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix}.$$

It is easy to verify that M and D are positive definite, and $\|D\| \leq 1$. However, the matrix

$$DM + I - D = \begin{pmatrix} 6 & 2 \\ -1 & 0 \end{pmatrix}$$

cannot be strongly monotone since it has a zero diagonal entry.

Functions over Euclidean spaces

We now specialize Definition 3.2.1 for functions defined over an Euclidean space.

Property 3.2.3. ([8]) There is a constant $\alpha > 0$ such that for any $d_1, \ldots, d_r \geq 0$, any $d_{ij} \geq 0$, any Jordan frame $\{c_1, \ldots, c_r\}$ and any $x, y \in \mathbb{E}$,

$$\left\| F(x) - F(y) + \sum d_i(x_i - y_i)c_i + \sum d_{ij}(x_{ij} - y_{ij}) \right\| \geq \alpha \|x - y\|,$$

where $x = \sum_{i=1}^r x_i c_i + \sum_{1 \leq i < j \leq r} x_{ij}$ and $y = \sum_{i=1}^r y_i c_i + \sum_{1 \leq i < j \leq r} y_{ij}$ are Peirce decompositions.

It is not difficult to see that Property 3.2.3 is a special case of Definition 3.2.1 with the set of linear transformations defined as

$$\mathcal{L} := \{x \mapsto D \bullet x : \ D = [d_{ij}] \in \mathbb{S}^r, \ d_{ij} \geq 0\}$$

where $D \bullet x := \sum_{i=1}^r d_{ii} x_i c_i + \sum_{1 \leq i < j \leq r} d_{ij} x_{ij}$ and $x = \sum_{i=1}^r x_i c_i + \sum_{1 \leq i < j \leq j} x_{ij}$ is the Peirce decomposition of x with respect to the Jordan frame $\{c_1, c_2, ..., c_r\}$.

For the discussions below, we list several existing P-type properties in the literature.

Definition 3.2.4. A transformation $F : \mathbb{E} \to \mathbb{E}$ is said to satisfy

- the *uniform* Cartesian P-property if there exists $\rho > 0$ such that for any $x, y \in \mathbb{E}$,

$$\max_{1 \le v \le \kappa} \langle (x-y)_v, (F(x) - F(y))_v \rangle \ge \rho \|x - y\|^2$$

  where x_ν denotes the ν-th component of x in the direct sum $\mathbb{E} = \mathbb{E}_1 \oplus \cdots \oplus \mathbb{E}_\kappa$ of Euclidean Jordan algebras.

- the *uniform* Jordan P-property [59] if there exists $\alpha > 0$ such that for any $x, y \in \mathbb{E}$,

$$\lambda_1((x-y) \circ (F(x) - F(y))) \ge \alpha \|x - y\|^2.$$

- the *uniform* P-property [59] if there exists $\alpha > 0$ such that for any $x, y \in \mathbb{E}$ with $x - y$ operator commuting with $F(x) - F(y)$,

$$\lambda_1((x-y) \circ (F(x) - F(y))) \ge \alpha \|x - y\|^2.$$

Remark 3.2.1. The uniform Cartesian P-property is a straightforward extension of the one introduced by Chen and Qi in [3]. When F is linear, i.e., $F(x) = L(x) + q$, the uniform Cartesian P-property reduces to the *Cartesian P-property*:

$$\max_{1 \le v \le \kappa} \langle L(x)_v, x_v \rangle > 0, \ \forall \ x \ne 0.$$

Remark 3.2.2. When F is linear, the uniform Jordan P-property reduces to the *Jordan P-property* [57]:

$$(x-y) \circ (F(x) - F(y)) \le_{\mathcal{K}} 0 \implies x = y,$$

and the uniform P-property reduces to the *P-property* [57]:

$$\left.\begin{array}{r} x-y \text{ and } F(x)-F(y) \text{ operator commute} \\ (x-y) \circ (F(x) - F(y)) \le_{\mathcal{K}} 0 \end{array}\right\} \implies x = y.$$

We now make connections to the above properties.

Proposition 3.2.4 ([8])**.** *Let $\mathbb{E} = \mathbb{E}_1 \oplus \cdots \oplus \mathbb{E}_\kappa$. If $F : \mathbb{E} \to \mathbb{E}$ has the uniform Cartesian P-property, then it has the Property 3.2.3.*

A weaker version of the uniform Cartesian P-property is the following: there exists $\gamma > 0$ such that for any $x, y \in \mathbb{E}$,

$$\max_{1 \le v \le \kappa} \frac{\langle (x-y)_v, (F(x) - F(y))_v \rangle}{\|(x-y)_v\|} \ge \gamma \|x - y\|. \tag{3.7}$$

The normalization above is interpreted as 0 when $x_v = y_v$. Clearly, if $\kappa = 1$, the uniform Cartesian P-property reduces to the strong monotonicity, and hence the uniform nonsingularity over the set of self-adjoint monotone linear transformations. When $\kappa \ne 1$, we show that (3.7) is equivalent to another special case of Definition 3.2.1.

Given $\mathbb{E} = \mathbb{E}_1 \oplus \cdots \oplus \mathbb{E}_\kappa$, we consider a linear transformations D over $\mathbb{E}$ of the form

$$D = D_1 \oplus \cdots \oplus D_\kappa$$

such that

$$(Dx)_\nu = D_\nu x_\nu \quad \forall\ 1 \le \nu \le \kappa,$$

where D_ν is a linear transformation over $\mathbb{E}_\nu$. It is clear that if D is self-adjoint and monotone over $\mathbb{E}$ if and only if each D_ν is self-adjoint and monotone over $\mathbb{E}_\nu$.

Property 3.2.4. There is a constant $\alpha > 0$ such that for any self-adjoint monotone linear transformation $D : \mathbb{E} \to \mathbb{E}$ and for any $x, y \in \mathbb{E}$

$$\|F(x) - F(y) + D(x - y)\| \ge \alpha \|x - y\|,$$

where $\mathbb{E} = \mathbb{E}_1 \oplus \cdots \oplus \mathbb{E}_\kappa$ and $D = D_1 \oplus \cdots \oplus D_\kappa$.

Proposition 3.2.5. *The transformation $F : \mathbb{E} \to \mathbb{E}$ satisfies* (3.7) *if and only if it has the Property 3.2.4.*

Proof. "Only if": Assume that (3.7) is satisfied. Let $D = D_1 \oplus \cdots \oplus D_\kappa$ be an self-adjoint monotone linear transformation over $\mathbb{E}$. It straightforwardly follows from the Cauchy-Schwartz inequality and monotonicity of D that

$$\begin{aligned}
\|F(x) - F(y) + D(x - y)\| &\ge \max_{1 \le v \le \kappa} \|(F(x) - F(y))_v + D_v(x - y)_v\| \\
&\ge \max_{1 \le v \le \kappa} \frac{\langle (F(x) - F(y))_v + D_v(x - y)_v, (x - y)_v \rangle}{\|(x - y)_v\|} \\
&\ge \max_{1 \le v \le \kappa} \frac{\langle (x - y)_v, (F(x) - F(y))_v \rangle}{\|(x - y)_v\|} \ge \gamma \|x - y\|.
\end{aligned}$$

"If": Fix any $x, y \in \mathbb{E}$. By the argument used in the proof of the if part of Proposition 3.2.3, we have

$$\begin{aligned}
&\inf \|F(x) - F(y) + D(x - y)\|^2 \\
&= \sum_{v=1}^{\kappa} \inf \|(F(x) - F(y))_v + D_v(x - y)_v\|^2 \\
&= \sum_{v \in I_+} \frac{\langle (x - y)_v, (F(x - F(y))_v \rangle^2}{\|(x - y)_v\|^2} + \sum_{v \in I_0} \|F(x)_v - F(y)_v\|^2,
\end{aligned}$$

where the index sets

$$I_+ = \{v : \langle (x - y)_v, (F(x) - F(y))_v \rangle > 0\},$$

and

$$I_0 = \{v : \langle (x - y)_v, (F(x) - F(y))_v \rangle = 0,\ x_v - y_v = 0 \text{ and } F(x)_v \ne F(y)_v\}.$$

Foe the ease of simplicity, we also denote by

$$I_- = \{v : \langle (x-y)_v, (F(x)-F(y))_v \rangle < 0\}.$$

Consider the perturbation $x \mapsto x - \varepsilon\Delta$, where $\varepsilon > 0$ and $\Delta = F(x) - F(y)$. When ε is sufficiently small, it follows from the continuity that:

(1) $\langle (x-\varepsilon\Delta - y)_v, (F(x-\varepsilon\Delta)-F(y))_v \rangle > 0$ if $v \in I_+$;

(2) $\langle (x-\varepsilon\Delta - y)_v, (F(x-\varepsilon\Delta)-F(y))_v \rangle < 0$ if $v \in I_0$;

(3) $\langle (x-\varepsilon\Delta - y)_v, (F(x-\varepsilon\Delta)-F(y))_v \rangle < 0$ if $v \in I_-$;

(4) $\|(F(x-\varepsilon\Delta x)-F(y))_v + D_v(x-\varepsilon\Delta - y)_v\| \to 0$ as $\varepsilon \to 0$ if $v \notin I_+ \cup I_0 \cup I_-$.

Consequently, for ε sufficiently small, we have

$$\begin{aligned}
&\alpha^2\|x-\varepsilon\Delta - y\|^2 \\
&\le \inf\|F(x-\varepsilon\Delta)-F(y)+D(x-\varepsilon\Delta x - y)\|^2 \\
&= \sum_{v\in I_+} \frac{\langle (x-\varepsilon\Delta - y)_v, (F(x-\varepsilon\Delta)-F(y))_v \rangle^2}{\|(x-\varepsilon\Delta-y)_v\|^2} + o(\varepsilon),
\end{aligned}$$

where the term $o(\varepsilon) \to 0$ as $\varepsilon \downarrow 0$. Taking limit $\varepsilon \downarrow 0$, we get

$$\sum_{v\in I_+} \frac{\langle (x-y)_v, (F(x-F(y))_v \rangle^2}{\|(x-y)_v^2\|^2} \ge \alpha^2\|x-y\|^2.$$

Therefore, there exists some index v such that

$$\frac{\langle (x-y)_v, (F(x-F(y))_v \rangle}{\|(x-y)_v\|} \ge \frac{\alpha}{\sqrt{\kappa}}\|x-y\|,$$

which shows that (3.7) is satisfied. □

It was shown in [8] that, when F is linear, then the Property 3.2.3 implies the uniform P-property. However, it is not clear how they are related in general. The next result shows that these properties are equivalent for *Löwner's operators*. Given a scalar valued function $\phi : \mathbb{R} \to \mathbb{R}$, the corresponding *Löwner's operator* is defined by

$$x \in \mathbb{E} \mapsto \sum_{i=1}^{r} \phi(\lambda_i(x))c_i,$$

where $x = \sum_{i=1}^{r} \lambda_i(x)c_i$ is any spectral decomposition. We shall denote it by $\phi_{\mathbb{E}}$.

Theorem 3.2.1. *For the Löwner's operator $\phi_{\mathbb{E}}$, the following are equivalent.*

(1) ϕ is strongly increasing;

(2) $\phi_{\mathbb{E}}$ has strong monotonicity property.

(3) $\phi_{\mathbb{E}}$ has the uniform Cartesian P-property;

(4) $\phi_{\mathbb{E}}$ has the Property 3.2.3;

(5) $\phi_{\mathbb{E}}$ has the uniform Jordan P-property;

(6) $\phi_{\mathbb{E}}$ has the uniform P-property.

Proof. *(1) $\Longrightarrow$ (2)*: It follows from Theorem 9 in [28].

(2) $\Longrightarrow$ (3): In general, strong monotonicity implies the uniform Cartesian P-property as

$$\begin{aligned}\langle x-y, F(x)-F(y)\rangle &= \sum_{v=1}^{\kappa}\langle (x-y)_v, (F(x)-F(y))_v\rangle \\ &\leq \kappa \max_{1\leq v\leq\kappa}\langle (x-y)_v, (F(x)-F(y))_v\rangle .\end{aligned}$$

(3) $\Longrightarrow$ (4): It follows from Theorem 3.2.4.

(4) $\Longrightarrow$ (1): By using multiples of the unit e and the choices $d_i = d$ and $d_{ij} = 0$ for all i, j in the definition of uniform nonsingularity, we deduce that for any $d \geq 0$ and any $x > y$,

$$|\phi(x) - \phi(y) + d(x-y)|\|e\| = \|\phi_{\mathbb{E}}(xe) - \phi_{\mathbb{E}}(ye) + d(x-y)e\| \geq \alpha(x-y)\|e\|.$$

Hence $\phi(x) \geq \phi(y)$ for any $x > y$. Moreover, with $d = 0$, we get $\phi(x)-\phi(y) = |\phi(x)-\phi(y)| \geq \alpha(x-y)$; i.e., ϕ is strongly increasing.

The implications (2) $\Longrightarrow$ (5) $\Longrightarrow$ (6) follows from [59, Proposition 3.1].

(6) $\Longrightarrow$ (1): By using multiples of the unit e in the definition of the uniform P-property, we deduce that for any $x > y$,

$$(\phi(x) - \phi(y))(x-y) = \lambda_1((\phi_{\mathbb{E}}(xe) - \phi_{\mathbb{E}}(ye)) \circ (x-y)e) \geq \alpha(x-y)^2,$$

completing the proof. □

Next, we consider another special transformation, namely the Lyapunov transformation. Recall that for a given $a \in \mathbb{E}$, the Lyapunov transformation is defined by

$$L_a(x) = a \circ x, \ \forall\, x \in \mathbb{E}.$$

Theorem 3.2.2. *For the Lyapunov transformation L_a, the following are equivalent:*

(1) $a \in int(\mathcal{K})$;

(2) L_a has strict monotonicity property;

(3) L_a *has the Cartesian* P*-property;*

(4) L_a *has the Property 3.2.3;*

(5) L_a *has the Jordan* P*-property;*

(6) L_a *has the* P*-property.*

Proof. *(1)* $\Longrightarrow$ *(2)*: It follows from Theorem 1.3.2.

(2) $\Longrightarrow$ *(3)*: See proof of Theorem 3.2.1.

(3) $\Longrightarrow$ *(4)*: See Proposition 4.5 in [8].

(4) $\Longrightarrow$ *(6)*: It follows from [8, Proposition 4.4].

(6) $\Longrightarrow$ *(1)*: Let $a = \sum \lambda_i(a)c_i$ be a spectral decomposition. Let I_- be the index set $\{i : \lambda_i(a) \leq 0\}$, which is empty if and only if $a \in \mathrm{int}(\mathcal{K})$; see Theorem 1.3.2. Consider $x = \sum_{i \in I_-} c_i$, which is zero if and only if I_- is empty. Noting that x and $L_a(x)$ operator commute and

$$x \circ L_a(x) = \sum_{i \in I_-} \lambda_i(a)c_i \leq_{\mathcal{K}} 0,$$

we deduce that if L_a satisfies the P-property, then x must be zero, which eventually leads to $a \in \mathrm{int}(\mathcal{K})$.

The implication (2) $\Longrightarrow$ (5) $\Longrightarrow$ (6) follows from [57, Theorem 11]. □

We next introduce another special case of (3.2), which involves the notion of the generalized Jacobian, see, e.g., [9]. It is well known that $\Pi_{\mathcal{K}}$ is nonexpansive and globally Lipschitz continuous, hence the Clarke's generalized Jacobian is well defined everywhere by Rademacher's theorem. Let $\mathcal{D}_\Pi$ be the set of points at which $\Pi_{\mathcal{K}}$ is differentiable. The Clarke's generalized Jacobian at x, denoted $\partial\Pi_{\mathcal{K}}(x)$, is defined by

$$\partial\Pi_{\mathcal{K}}(x) = \mathrm{conv}\{\lim_{k\to\infty} \nabla\Pi_{\mathcal{K}}(x^k) : x^k \to x, x^k \in \mathcal{D}_\Pi\}.$$

It can be shown that the set $\partial\Pi_{\mathcal{K}}(x)$ is convex and compact (see, e.g., [9]). We list below some useful properties of the generalized Jacobian.

Proposition 3.2.6. *The following statements are true*

(a) Every element D *of* $\partial\Pi_{\mathcal{K}}(x)$ *is self-adjoint and monotone with* $\|D\| \leq 1$.

(b) For any $x, y \in \mathbb{E}$*, there exists some* $D \in \mathrm{conv}\{\partial\Pi_{\mathcal{K}}([x, y])\}$ *such that*

$$\Pi_{\mathcal{K}}(x) - \Pi_{\mathcal{K}}(y) = D(x - y).$$

(c) $\partial\Pi_{\mathcal{K}}(x) \subseteq \partial\Pi_{\mathcal{K}}(0)$ *for all* $x \in \mathbb{E}$.

Proof. (a). See the proof of Theorem 2.2 in [53].

(b). See the proof of Proposition 2.6.5 in [9].

(c). Fix any $y \in \mathcal{D}_\Pi$. Since $\Pi_\mathcal{K}$ is positively homogeneous, it follows that

$$\nabla\Pi_\mathcal{K}(ty) = \nabla\Pi_\mathcal{K}(y)$$

for any $t > 0$. Letting $t \downarrow 0$, we conclude that $\nabla\Pi_\mathcal{K}(y) \in \partial\Pi_\mathcal{K}(0)$. This, together with convexity and compactness of $\partial\Pi_\mathcal{K}(0)$, has in fact proved that $\partial\Pi_\mathcal{K}(x) \subseteq \partial\Pi_\mathcal{K}(0)$. □

When $\mathcal{K} = \mathbb{R}^n_+$, it can be verified that $\partial\Pi_\mathcal{K}(0)$ consists of diagonal matrices D with $0 \leq D \leq I$.

The following result follows exactly from the argument used in the proof of Lemma (3.2.1).

Proposition 3.2.7. *Let $L : \mathbb{E} \to \mathbb{E}$ be a linear transformation, and let $K \subset \mathbb{E}$ be a closed convex cone. The following statements are equivalent:*

(1) There exists $\alpha > 0$ such that

$$\|(LD + I - D)v\| \geq \alpha\|v\|,$$

for all $v \in \mathbb{E}$ and $D \in \partial\Pi_\mathcal{K}(0)$.

(2) There exists $\beta > 0$ such that

$$\left\|(L + D^{-1} - I)v\right\| \geq \beta\|v\|,$$

for all $v \in \mathbb{E}$ and all nonsingular $D \in \partial\Pi_\mathcal{K}(0)$.

(3) There exists $\gamma > 0$ such that

$$\|(DL + I - D)v\| \geq \gamma\|v\|,$$

for all $v \in \mathbb{E}$ and all $D \in \partial\Pi_\mathcal{K}(0)$.

For a linear transformation $L : \mathbb{E} \to \mathbb{E}$, recall from Chapter 1 that the normal map is defined by

$$F^{\mathrm{nor}}(x) = L(\Pi_\mathcal{K}(x)) + x - \Pi_\mathcal{K}(x).$$

Adapting the argument in the proof of Proposition 1.5.11 in [11], we know that L has the globally unique solvability (GUS) property if and only if F is a homeomorphism of $\mathbb{E}$.

Definition 3.2.5. ([57]) A linear transformation $L : \mathbb{E} \to \mathbb{E}$ is said to have the Lipschitzian GUS-property if the normal map $F^{\mathrm{nor}}(z) = L(\Pi_\mathcal{K}(z)) + z - \Pi_\mathcal{K}(z)$ is a homeomorphism of $\mathbb{E}$ and the inverse of F^{nor} is Lipschitzian.

The above property holds [57] for L if and only if there exists a β such that

$$\|F^{\text{nor}}(x) - F^{\text{nor}}(y)\| \geq \beta\|x - y\|$$

for all $x, y \in \mathbb{E}$. Furthermore, it is well known [11] that the above property holds if and only if the the solution map of the $\mathbf{CP}(L, \mathcal{K})$ is a homeomorphism and a Lipschitz function.

Gowda et al [57] described some necessary conditions for the Lipschizian GUS-property of L, which is also sufficient when $\mathcal{K}$ is polyhedral. Our next result give a sufficient condition for the Lipschizian GUS-property of L.

Proposition 3.2.8. *If L is uniformly nonsingular with modulus α over $\mathcal{L} = \{D^{-1} - I : D \in \partial\Pi_{\mathcal{K}}(0),\ D$ nonsingular$\}$, then it has the Lipschitzian GUS-property.*

Proof. Note that $\Pi_{\mathcal{K}}(x) - \Pi_{\mathcal{K}}(y) = D(x - y)$ for some $D \in \partial\Pi_{\mathcal{K}}(0)$. Thus, by the uniform nonsingularity of L and Proposition 3.2.7, we have

$$\begin{aligned}\|F^{\text{nor}}(x) - F^{\text{nor}}(y)\| &= \|L(\Pi_{\mathcal{K}}(x) - \Pi_{\mathcal{K}}(y)) + (x - y) + (\Pi_{\mathcal{K}}(y) - \Pi_{\mathcal{K}}(x))\| \\ &= \|(LD + I - D)(x - y)\| \\ &\geq \frac{\alpha}{\alpha + 1 + \|L\|}\|x - y\|,\end{aligned}$$

as desired. □

Next, we establish an error bound for the $\mathbf{CP}(F, \mathcal{K})$, where $F(x) = L(x) + q$. Let $r(x) = \|x - \Pi_{\mathcal{K}}(x - L(x) - q)\|$. Clearly, if x^* is the solution of the $\mathbf{CP}(F, \mathcal{K})$ then $r(x^*) = 0$.

Proposition 3.2.9. *Suppose that $F(x) = L(x) + q$ and F is uniformly nonsingular over $\mathcal{L} = \{D^{-1} - I : D \in \partial_{\Pi_{\mathcal{K}}}(0), D$ nonsingular$\}$ with modulus α. Let x^* be the unique solution of the **CP**$(F, \mathcal{K})$. Then, for all $x \in \mathbb{E}$, it holds*

$$\frac{1}{1 + \|L\|}r(x) \leq \|x - x^*\| \leq \frac{\alpha + 1 + \|L\|}{\alpha}r(x).$$

Proof. Note that $\Pi_{\mathcal{K}}(x) - \Pi_{\mathcal{K}}(y) = D(x - y)$ for some $D \in \partial\Pi_{\mathcal{K}}(0)$. Using the uniform nonsingularity of L and Proposition 3.2.7, we obtain

$$\begin{aligned}r(x) &= \|[x - \Pi_{\mathcal{K}}(x - L(x) - q)] - [x^* - \Pi_{\mathcal{K}}(x^* - L(x^*) - q)]\| \\ &= \|(DL + I - D)(x - x^*)\| \\ &\geq \frac{\alpha}{\alpha + 1 + \|L\|}\|x - x^*\|.\end{aligned}$$

Thus, $\|x - x^*\| \leq \frac{\alpha+1+\|L\|}{\alpha}r(x)$.

On the other hand, since $\|D\| \leq 1$, we have

$$\begin{aligned}r(x) &= \|(DL + I - D)(x - x^*)\| \leq (\|DL\| + \|I - D\|)\,\|x - x^*\| \\ &\leq (1 + \|L\|)\|x - x^*\|,\end{aligned}$$

implying that $\frac{1}{1+\|L\|}r(x) \leq \|x - x^*\|$. □

3.3 Conclusion

In this chapter, we have proposed the general notion of uniform nonsingularity property for transformations over Euclidean spaces, which is closely related to a number of existing properties. In particular, it recovers the P-type property (3.2.3) studied in [8] which represents a class of numerically solvable non-monotone SCCPs. A variant of the property yields a new characterization of the strong monotonicity, and some form of the property of a linear transformation implies the Lipschitzian GUS-property. In next chapter, we shall present the application of the uniform nonsingularity property for the $\mathbf{CP}(F, \mathcal{K})$ over a general closed convex cone.

Chapter 4

A Smoothing Homotopy Path to The Unique Solution

4.1 Introduction

In Chapter 3, we proposed the notion of uniform non-singularity property (3.2) for transformations over Euclidean spaces and discussed its relationships with several existing properties. In this chapter, we aim to investigate its connection with the uniqueness of solution of the $\mathbf{CP}(F, \mathcal{K})$ (cf. (1.7)).

Our analysis is based on one of the most well-known reformulations for complementarity problems - the normal map equation (NME):

$$F(\Pi_{\mathcal{K}}(z)) + z - \Pi_{\mathcal{K}}(z) = 0. \tag{4.1}$$

In this approach, every solution to the $\mathbf{CP}(F, \mathcal{K})$ corresponds exactly to a solution to the NME via $x = \Pi_{\mathcal{K}}(z)$ and $z = x - F(x)$. The main difficulty in solving the NME comes from the nondifferentiability of $\Pi_{\mathcal{K}}$. Among various methods proposed to overcome this difficulty is the use of smoothing approximations of $\Pi_{\mathcal{K}}$. The basis for this approach is to construct a continuously differentiable function $G : \mathbb{E} \times \mathbb{R}_{++} \to \mathbb{E}$, parameterized by $\mu \in \mathbb{R}_{++}$, such that G approaches the projection function $\Pi_{\mathcal{K}}$ as μ approaches zero, i.e.,

$$\lim_{\mu \downarrow 0} G(x, \mu) = \Pi_{\mathcal{K}}(x) \quad \forall x \in \mathbb{E}.$$

For convenience, we denote $G(x, 0) = \Pi_{\mathcal{K}}(x)$.

In this chapter we study a barrier-based method [7] to construct smoothing approximations of $\Pi_{\mathcal{K}}$ of an arbitrary closed convex cone $\mathcal{K}$, whose derivatives depend on those of the barrier used. Accordingly, (4.1) can be approximated by the following parametric equation, called the smooth normal map equation (SNME):

$$(1 - \mu)F(G(z, \mu)) + z - (1 - \mu)G(z, \mu) - \mu b = 0, \tag{4.2}$$

where $b \in \mathbb{E}$ is fixed, $\mu \in (0, 1]$. When $\mu = 1$, the SNME becomes

$$z - b = 0,$$

which has the uniques solution $z = b$. On the other hand, when $\mu = 0$, the SNME reduces to the NME. Therefore, if there exists a trajectory from the unique solution at $\mu = 1$ to a solution at $\mu = 0$, we can apply standard homotopy techniques to find the solution of the NME, and hence a solution of $\mathbf{CP}(F, \mathcal{K})$. In Section 4.3, we shall show that the uniform nonsingularity property ensures the existence and boundedness of a convergent trajectory. Moreover, we show that every solution of the $\mathbf{CP}(F, \mathcal{K})$ is the limit of the trajectory, whence establishing the uniqueness of the solution of $\mathbf{CP}(F, \mathcal{K})$.

4.2 Barrier Based Smoothing Approximation

Let $f : \text{int}(\mathcal{K}) \to \mathbb{R}$ be a strictly convex twice differentiable barrier function. Being a barrier means that $f(x_k) \to \infty$ for any sequence $\{x_k\} \subset \text{int}(\mathcal{K})$ converging to the boundary $bd(\mathcal{K})$. Thus, for any sequence $\{x_k\} \subset \text{int}(\mathcal{K})$ converging to the boundary $bd(\mathcal{K})$, it follows from the convexity of f that $\|\nabla f(x_k)\| \|x_k - x_1\| \geq \langle \nabla f(x_k), x_k - x_1 \rangle \geq f(x_k) - f(x_1) \to \infty$, whence $\|\nabla f(x_k)\| \to \infty$.

To derive interesting properties, we further assume that f satisfies the following assumptions.

(A1) For any $x \in \text{int}(\mathcal{K})$ and any $t \in \mathbb{R}_{++}$, it holds that

$$\nabla f(tx) = \frac{1}{t} \nabla f(x).$$

(A2) The following bound holds

$$\sup_{x \in \text{int}(\mathcal{K})} \langle x, -\nabla f(x) \rangle < \infty. \tag{4.3}$$

Note that differentiating the equation in Assumption A(1) with respect to x gives

$$\nabla^2 f(tx) = \frac{1}{t^2} \nabla^2 f(x).$$

These assumptions are satisfied by logarithmically homogeneous barriers; i.e., barriers f satisfying

$$\forall x \in \text{int}(\mathcal{K}),\ \forall t \in \mathbb{R}_{++},\ f(tx) = f(x) - \vartheta \log t$$

for some $\vartheta \geq 0$.

We shall show that for each $\mu > 0$, the function $\varrho_\mu : x \mapsto x + \mu \nabla f(x) = x + \nabla f(\frac{x}{\mu})$ is a bijection between $\text{int}(\mathcal{K})$ and $\mathbb{E}$, and that its inverse function is a smoothing approximation of $\Pi_{\mathcal{K}}$. To achieve that, we first introduce the definition and basic results of *maximal monotone* set-valued map $T : \mathbb{H} \rightrightarrows \mathbb{H}$ over a Hilbert space $\mathbb{H}$. For a more comprehensive introduction, we refer to Rockafellar's classic papers [48] and [47].

Definition 4.2.1. A set-valued map $T : \mathbb{H} \rightrightarrows \mathbb{H}$ over a Hilbert space $\mathbb{H}$ with inner product $\langle \cdot, \cdot \rangle$ is said to be a *monotone operator* if

$$\forall (z, w), (z', w') \in \mathcal{G}(T), \quad \langle z - z', w - z' \rangle \geq 0,$$

where $\mathcal{G}(T) \triangleq \{(z, w) \in \mathbb{H} \times \mathbb{H} : w \in T(z)\}$ denotes the *graph* of T. A monotone operator is *maximal monotone* if its graph is not contained in the graph of another monotone operator. Equivalently, a monotone operator is maximal operator if

$$\forall (z, w) \in \mathbb{H} \times \mathbb{H}, \left(\inf_{(z', w') \in \mathcal{G}(T)} \langle z - z', w - z' \rangle \geq 0 \Longrightarrow (z, w) \in \mathcal{G}(T) \right).$$

It is easy to see that (maximal) monotonicity is preserved by positive scalings, that is, T is (maximal) monotone if and only if cT is (maximal) monotone, where $c > 0$. We shall make use of the following two characterizations of the maximal monotonicity.

Theorem 4.2.1 (Minty's characterization of maximal monotonicity [38])**.** *For each $\lambda > 0$, a monotone set-valued map $M : \mathbb{E} \rightrightarrows \mathbb{E}$ is maximally monotone if and only if $I + \lambda M$ is a bijection between dom(M) and $\mathbb{E}$.*

Theorem 4.2.2 (Löhne's characterization of maximal monotonicity [31])**.** *A set-valued map $M : \mathbb{E} \rightrightarrows \mathbb{E}$ is maximally monotone if and only if the following are satisfied.*

(i) M is monotone.

(ii) M has a nearly convex domain (i.e., cl(dom(M)) is convex).

(iii) The values of M are convex.

(iv) The recession cone of $M(x)$ equals the normal cone to cl(dom(M)) at every $x \in$ dom(M).

(v) The graph of M is closed.

Proposition 4.2.1. *For each $\mu > 0$, the function ϱ_μ is a bijection between int$(\mathcal{K})$ and $\mathbb{E}$.*

Proof. We shall use Löhne's characterization to check that the gradient map $x \mapsto \{s : s = \nabla f(x)\}$ is maximally monotone, whence deduce the proposition by Minty's characterization. Indeed,

(i) the gradient map is the derivative map of a convex function, whence monotone;

(ii) the domain $\mathrm{int}(\mathcal{K})$ of the gradient map is convex;

(iii) the values of ∇f are singletons, whence convex;

(iv) the recession cone of the gradient map at each x in its domain $\mathrm{int}(K)$ is the trivial subspace $\{0\}$, which agrees with the normal cone to $cl(\mathrm{int}(\mathcal{K})) = \mathcal{K}$ at x;

(v) the graph of the gradient map is closed as ∇f is continuous and $\|\nabla f(x_k)\| \to \infty$ for any sequence $\{x_k\} \subset \text{int}(\mathcal{K})$ converging to the boundary $bd(\mathcal{K})$.

□

Proposition 4.2.2. *The inverse functions p_μ of ϱ_μ for $\mu > 0$ are continuously differentiable functions with Jacobian $Jp_\mu(z) = (I + \mu\nabla^2 f(p_\mu(z)))^{-1} = (I + \nabla^2 f(\frac{1}{\sqrt{\mu}}p_\mu(z)))^{-1}$, and satisfy the point-wise convergence $p_\mu \overset{\mu\to 0}{\longrightarrow} \Pi_{\mathcal{K}}$.*

Proof. The continuous differentiability of p_μ and its Jacobian follows from the Inverse Function Theorem as $J\varrho_\mu(x) = I + \mu\nabla^2 f(x)$ is nonsingular for each $x \in \text{int}(\mathcal{K})$.

To prove point-wise convergence we first establish that $\{p_\mu(z) : \mu \in (0, 1)\}$ is bounded for each $z \in \mathbb{E}$, then deduce that every convergent sequence $\{p_{\mu_k}(z)\}$ with $\mu_k \to 0$ and $z \in \mathbb{E}$ has the limit $\Pi_{\mathcal{K}}(z)$.

Fix an arbitrary $e \in \text{int}(\mathcal{K})$ and consider the bounded set $\{\varrho_\mu(e) = e + \mu\nabla f(e) : \mu \in (0, 1)\}$. It was established by Rockafellar [47] that p_μ is nonexpansive for each $\mu > 0$. Thus, for each $z \in \mathbb{E}$, $\|p_\mu(z)\| \le \|p_\mu(z) - p_\mu(\varrho_\mu(e))\| + \|e\| \le \|z - \varrho_\mu(e)\| + \|e\|$ for all $\mu > 0$ shows that $\{p_\mu(z) : \mu \in (0, 1)\}$ is bounded.

For each $z \in \mathbb{E}$, every sequence $\mu_k \to 0$, and every convergent $\{p_{\mu_k}(z)\}$, say with limit x_∞, we have $\text{int}(\mathcal{K}) \ni p_{\mu_k}(z) \to x_\infty$, $\text{int}(\mathcal{K}^*) \ni -\mu_k\nabla f(p_{\mu_k}(z)) = p_{\mu_k}(z) - z \to x_\infty - z$ and

$$0 \le \langle p_{\mu_k}(z), -\mu_k\nabla f(p_{\mu_k}(z))\rangle \le \mu_k \sup_{x\in\text{int}(\mathcal{K})} \langle x, -\nabla f(x)\rangle \overset{(4.3)}{\to} 0,$$

whence $x_\infty \in \mathcal{K}$, $x_\infty - z \in \mathcal{K}^*$ and $\langle x_\infty, x_\infty - z\rangle = 0$. Hence $x_\infty = \Pi_{\mathcal{K}}(z)$ as required. □

When the underlying cone $\mathcal{K}$ is symmetric, the barried-based smoothing approximations simplify to the classic smoothing approximations, see, e.g., [1, 27, 50].

Example 4.2.1. *Consider $\mathcal{K} = \mathbb{R}^n_+$, the cone of nonnegative orthant. A barrier f of $\mathcal{K}$ that satisfies all assumptions is $f : x \in \mathbb{R}^n_{++} \mapsto -\log\prod_{i=1}^n x_i = -\sum_{i=1}^n \log x_i$. Simple calculations show that the function ϱ_μ is given by $x \in \mathbb{R}^n_{++} \mapsto x - \mu/x$, where $1/x$ denotes the vector with components $1/x_i$. Then the barrier-based smoothing approximation is $p_\mu : z \in \mathbb{R}^n \mapsto \frac{1}{2}(z + \sqrt{z^2 + 4\mu e})$, where e denotes the vector of all ones, $\sqrt{x}$ and x^2 denotes the vectors with components $\sqrt{x_i}$ and x_i^2, respectively. This coincides with the Chen-Harker-Kanzow-Smale (CHKS) smoothing approximation.*

Example 4.2.2. *Consider $\mathcal{K} = \mathbb{S}^n_+$, the cone of n-by-n real symmetric positive semidefinite matrices. A barrier f of $\mathcal{K}$ that satisfies all assumptions is the standard log-determinant barrier $f : X \in \mathbb{S}^n_{++} \mapsto -\log\det(X)$, where $\mathbb{S}^n_{++}$ denotes the set of symmetric positive definite matrices. For each $\mu > 0$, the function ϱ_μ is then $X \in \mathbb{S}^n_{++} \mapsto X - \mu X^{-1}$. It is straightforward to verify that its inverse function is $p_\mu : Z \in \mathbb{S}^n \mapsto \frac{1}{2}(Z + \sqrt{Z^2 + 4\mu I})$, where I denotes the n-by-n identity matrix, and $\sqrt{X}$ denotes the unique $Y \in \mathbb{S}^n_{++}$ such that $Y^2 = X \in \mathbb{S}^n_{++}$. This is also precisely the extension of the CHKS smoothing approximation.*

Example 4.2.3. *Let $\mathcal{K}$ be a symmetric cone of rank r, a r-logarithmatically homogeneous barrier function is $f : x \in int(\mathcal{K}) \mapsto -\log\det(x)$, where* det *is the determinant function in a Euclidean Jordan algebra $\mathbb{E}$. The smoothing approximation derived from the barrier function is $p_\mu : z \in \mathbb{E} \mapsto \frac{1}{2}(z + \sqrt{z^2 + 4\mu e})$, where e is the unitary element of $\mathbb{E}$ and $\sqrt{x}$ is the unique $y \in \mathfrak{J}$ such that $y^2 = x \in int(\mathcal{K})$. Again, this is another extension of the CHKS smoothing approximation.*

Compatability

We are interested in a class of smoothing functions which are compatible with a set $\mathcal{L}$ of linear operators in the following sense.

Definition 4.2.2. Let $\mathcal{L}$ be a set of linear transformations over $\mathbb{E}$. We say that the smoothing function G is *compatible with* $\mathcal{L}$ if the following two conditions are satisfied.

1. For any $\mu > 0$ and $x, v \in \mathbb{E}$, $J_x G(x, \mu)v = Dv$ for some $D \in \{(I + M)^{-1} : M \in \mathcal{L}\}$.
2. For any $\mu > 0, x \in \mathbb{E}$, $G(x, \mu) - G(0, \mu) = Dx$ for some $D \in \{(I + M)^{-1} : M \in \mathcal{L}\}$.

The expression for the Jacobian in Proposition 4.2.2 shows that the smoothing approximation $G : (x, \mu) \mapsto p_\mu(x)$ satisfies the first condition in the definition of compatibility with $\{\nabla^2 f(x) : x \in \text{int}(\mathcal{K})\}$. The following proposition then shows that the smoothing approximation G is compatible with

$$C = \{\nabla^2 f(x) : x \in \text{int}(\mathcal{K})\} \cup \left\{ \int_0^1 \nabla^2 f(tx + (1-t)e)\,\mathrm{d}t : x \in \text{int}(\mathcal{K}) \right\},$$

where $e = p_1(0)$ is the unique fixed point of $x \mapsto -\nabla f(x)$.

Proposition 4.2.3. *For each $\mu > 0$ and each $z \in \mathbb{E}$,*

$$p_\mu(z) - p_\mu(0) = \left(I + \int_0^1 \nabla^2 f\left(t\frac{p_\mu(z)}{\sqrt{\mu}} + (1-t)p_1(0) \right) \mathrm{d}t \right)^{-1} z.$$

Proof. Let e denote $p_1(0)$, which is the unique fixed point of $x \mapsto -\nabla f(x)$ by definition of p_1, and let x_t denote $tp_\mu(z) + (1-t)\sqrt{\mu}e$. It then follows from Assumptions A(1) and A(2) that $0 = \sqrt{\mu}e + \mu\nabla f(\sqrt{\mu}e) = \varrho_\mu(\sqrt{\mu}e)$; i.e., $p_\mu(0) = \sqrt{\mu}e$. Denoting $p_\mu(z)$ by x, we get $z = \varrho_\mu(x) = x + \mu\nabla f(x)$, and thus

$$\begin{aligned} p_\mu(z) - p_\mu(0) = x - \sqrt{\mu}e &= z - \mu(\nabla f(x) - \nabla f(\sqrt{\mu}e)) \\ &= z - \mu \int_0^1 \nabla^2 f(x_t)\,\mathrm{d}t(x - \sqrt{\mu}e) \\ &= z - \int_0^1 \nabla^2 f\left(\frac{x_t}{\sqrt{\mu}}\right) \mathrm{d}t(p_\mu(z) - p_\mu(0)). \end{aligned}$$

Consequently,

$$\left(I + \int_0^1 \nabla^2 f\left(\frac{x_t}{\sqrt{\mu}}\right) \mathrm{d}t\right)(p_\mu(z) - p_\mu(0)) = z.$$

This completes the proof. □

Example 4.2.4. *When $\mathcal{K} = \mathbb{S}^n_+$ and f is the standard log-determinant barrier, the Hessian $\nabla^2 f(x)$ is the linear map*

$$v \mapsto q((\lambda\lambda^T) \circ (q^T vq))q^T,$$

where λ is a vector of eigenvalues of x^{-1} and $x^{-1} = q\mathrm{Diag}(\lambda)q^T$ is a diagonalization. The unique fixed point $e = p_1(0)$ of $x \mapsto -\nabla f(x)$ is the identity matrix. Therefore, the set $\mathcal{L}$ mentioned above consists of linear transformations of the form $v \mapsto q(d \circ (q^T vq))q^T$, where q is any orthogonal matrix, and d either has rank one and positive entries or is a matrix of divided differences for $\log(\cdot)$. This set $\mathcal{L}$ is a subset of

$$\{v \mapsto q(d \circ (q^T vq))q^T : q \text{ orthogonal}, \ d \text{ symmetric with nonnegative entries}\}.$$

4.3 Globally Unique Solvability Property

In this section, we study the unique solvability of $\mathbf{CP}(F, \mathcal{K})$ under the following assumptions:

Assumption 4.3.1. 1. *The transformation $F : \mathbb{E} \to \mathbb{E}$ is continuously differentiable.*

2. *F is uniformly nonsingular (cf. (3.2)) over a set $\mathcal{L}$ of self-adjoint linear transformations on $\mathbb{E}$, where the $\mathcal{L}$ satisfies*

$$(1 + \lambda)(\mathcal{L} + \{I\}) \subseteq \mathcal{L} + \{I\} \quad \forall \lambda \geq 0.$$

3. *The smoothing approximation G of $\Pi_\mathcal{K}$ is compatible with $\mathcal{L}$ in the sense defined in Definition 4.2.2.*

Consider the map $H_b : \mathbb{E} \times \mathbb{R}_+ \to \mathbb{E}$ defined by

$$(z, \mu) \mapsto (1 - \mu)F(G(z, \mu)) + z - (1 - \mu)G(z, \mu) - \mu b, \tag{4.4}$$

where G is a smoothing approximation of Π_K, and $b \in \mathbb{E}$ is fixed. Recall that the SNME (4.2) is given by $H_b(z, \mu) = 0$. This homotopy is a slight modification of the one studied in [8] to describe and analyze a continuation method for solving SCCPs. Here the map is a homotopy between $z \mapsto z - b$ and $z \mapsto F(\Pi_\mathcal{K}(z)) + z - \Pi_\mathcal{K}(z)$.

Let S denote the set $\{(z, \mu) \in \mathbb{E} \times (0, 1] : H_b(z, \mu) = 0\}$. Define the solution path T as the connected component of S emanating from the unique solution of $H_b(z, 1) = 0$. In this section, we show that T forms a smooth and bounded trajectory that is monotone with respect to μ; i.e., there exists a continuously differentiable and bounded function $t \in I \mapsto (z(t), \mu(t))$ on some interval $I \subseteq \mathbb{R}$ such that $T = \{(z(t), \mu(t)) : t \in I\}$ and $t \mapsto \mu(t)$

is monotone. Thus, there exists at least a limit point as μ is reduced to zero along the trajectory. We further show that T is convergent and every solution of the NME (4.1) is the limit point of T.

The boundedness of the solution trajectory is studied first.

Lemma 4.3.1. *Let $\mathcal{S}_\delta(c)$ denote the level set*

$$\{(z,\mu) \in \mathfrak{J} \times (0, 1-\delta] : \|H_b(z,\mu)\| \leq c\}.$$

If Assumption 4.3.1 is satisfied, then for all c the level set $\mathcal{S}_\delta(c)$ is bounded for each $\delta > 0$.

Proof. For $\delta \in (0,1)$, and for arbitrary $(z,\mu) \in S_\delta(c)$, the compatibility of G with $\mathcal{L}$ implies $(I+M)(G(z,\mu) - G(0,\mu)) = z$ for some $M \in \mathcal{L}$. Therefore

$$\begin{aligned} &H_b(z,\mu) - (1-\mu)(F(G(0,\mu)) + G(0,\mu)) + \mu b \\ &= (1-\mu)\left\{F(G(z,\mu)) - F(G(0,\mu)) + \left(\frac{1}{1-\mu}(I+M) - I\right)(G(z,\mu) - G(0,\mu))\right\}. \end{aligned}$$

Together with the uniform nonsingularity of F over $\mathcal{L}$, say with modulus α, and the assumption $\bigcup_{\lambda \geq 0}(1+\lambda)(\mathcal{L} + \{I\}) \subseteq \mathcal{L} + \{I\}$, it follows

$$\|H_b(z,\mu) - (1-\mu)(F(G(0,\mu)) + G(0,\mu)) + \mu b\| \geq \alpha(1-\delta)\|G(z,\mu) - G(0,\mu)\|.$$

This gives an upper bound on $\|G(z,\mu)\|$ that is independent of z and μ. Consequently

$$\|z\| = \|H_b(z,\mu) - (1-\mu)(F(G(z,\mu)) - G(z,\mu)) + \mu b\|$$

has an upper bound independent of z and μ. This shows that the sub-level set $S_\delta(c)$ is bounded for $\delta \in (0,1)$. $\square$

The following proposition studies the nonsingularity of the Jacobian $J_z H_b(z,\mu)$.

Lemma 4.3.2. *Suppose Assumption 4.3.1 is satisfied, and $H_b(z,\mu)$ is as defined in* (4.4). *Then the Jacobian $J_z H_b(z,\mu)$ is nonsingular for any $(z,\mu) \in \mathbb{E} \times (0,1]$. Moreover, for each compact subset $U \subset \mathbb{E}$, there exists $\sigma > 0$ such that for all $(z,\mu) \in U \times (0, \frac{1}{2}]$ it holds that*

$$\|J_z H_b(z,\mu) v\| \geq \sigma \|v\|,$$

for any $v \in \mathbb{E}$.

Proof. For $\mu \in (0,1)$, we have

$$\begin{aligned} J_z H_b(z,\mu) v &= (1-\mu) JF(G(z,\mu)) J_z G(z,\mu) v + v - (1-\mu) J_z G(z,\mu) v \\ &= (1-\mu)\left\{JF(G(z,\mu)) w + \left(\frac{1}{1-\mu}(I+M) - I\right) w\right\}, \end{aligned}$$

where $w = J_zG(z,\mu)v = (I+M)^{-1}v$ for some $M \in \mathcal{L}$. Together with the uniform nonsingularity of F over $\mathcal{L}$, say with modulus α, and the assumption $\bigcup_{\lambda \geq 0}(1+\lambda)(\mathcal{L}+\{I\}) \subseteq \mathcal{L}+\{I\}$, it follows

$$\begin{aligned} \|J_zH_b(z,\mu)v\| &= (1-\mu)\left\|JF(G(z,\mu))w + \left(\frac{1}{1-\mu}(I+M) - I\right)w\right\| \\ &\geq \alpha(1-\mu)\|w\| \geq \frac{\alpha(1-\mu)}{\|I+M\|}\|v\|. \end{aligned} \tag{4.5}$$

Hence $J_zH_b(z,\mu)$ is nonsingular.

Let $U \subset \mathbb{E}$ be compact. The continuity of JF and G implies that

$$m_U \triangleq \sup\{\|JF(G(z,\mu))\| : (z,\mu) \in U \times [0, \frac{1}{2}]\} < \infty.$$

Then for $(z,\mu) \in U \times (0, \frac{1}{2}]$ and $v \in \mathbb{E}$, we have the following two cases.

1. If $\|J_zG(z,\mu)v\| < \frac{1}{(1-\mu)(\alpha+m_U+1)}\|v\|$, then

$$\begin{aligned} \|J_zH_b(z,\mu)v\| &= \|(1-\mu)JF(G(z,\mu))J_zG(z,\mu)v + v - (1-\mu)J_zG(z,\mu)v\| \\ &\geq \|v\| - (1-\mu)\|JF(G(z,\mu))J_zG(z,\mu)v\| - (1-\mu)\|J_zG(z,\mu)v\| \\ &\geq \|v\| - (1-\mu)(m_U+1)\|J_zG(z,\mu)v\| \\ &> \frac{\alpha}{\alpha+m_U+1}\|v\|; \end{aligned}$$

2. If $\|J_zG(z,\mu)v\| \geq \frac{1}{(1-\mu)(\alpha+m_U+1)}\|v\|$, then

$$\|J_zH_b(z,\mu)v\| \overset{(4.5)}{\geq} \alpha(1-\mu)\|J_zG(z,\mu)v\| \geq \frac{\alpha}{\alpha+m_U+1}\|v\|.$$

Therefore, for every compact $U \subset \mathbb{E}$,

$$\|J_zH_b(z,\mu)v\| \geq \frac{\alpha}{\alpha+m_U+1}\|v\| \tag{4.6}$$

for all $(z,\mu) \in U \times (0, \frac{1}{2}]$ and $v \in \mathbb{E}$. □

Now, by combining Lemma 4.3.1 and 4.3.2 and appealing to the Implicit Function Theorem (as in the proof of Proposition 5.4 in [8]), we can then establish the following result

Lemma 4.3.3. *If Assumption 4.3.1 is satisfied, then for each fixed $b \in \mathbb{E}$,*

(a) the SNME has a unique solution $z(\mu)$ for each $\mu \in (0,1]$;

(b) the set $T = \{(z_b(\mu),\mu) : \mu \in (0,1]\}$ is bounded;

(c) every accumulation point $(z^*, 0)$ *of* T *gives a solution* z^* *to the NME.*

Basically, the above result shows that $T = \{(z_b(\mu), \mu) : \mu \in (0, 1]\}$ is bounded and forms a smooth trajectory. Moreover, it has at least one accumulation point as the parameter μ decreases to zero along the trajectory, and every accumulation point is a solution of the NME. However, we have not ruled out the possibility of multiple accumulation points. We next show that T actually converges, and that every solution of the NME (and hence the $\mathbf{CP}(F, \mathcal{K})$) is the limit of the trajectory with $b = e$, whence establishing the uniqueness of the solution of $\mathbf{CP}(F, \mathcal{K})$.

Theorem 4.3.1. *If Assumption 4.3.1 holds and* $B \triangleq \sup\{\|\nabla_\mu G(z_b, \mu)\| : (z_b, \mu) \in T\} < \infty$, *then the* $\boldsymbol{CP}(F, \mathcal{K})$ *has a unique solution.*

Proof. We have already established the existence of solution. To prove uniqueness, we shall show that $z_b(\mu)$ converges as $\mu \downarrow 0$; and that every solution to $\mathbf{CP}(F, \mathcal{K})$ is an accumulation point of $z_0(\mu)$ when $\mu \downarrow 0$.

Differentiating $H_b(z_b(\mu), \mu) = 0$ with respect to μ gives

$$\nabla_\mu H_b(z_b(\mu), \mu) + J_z H_b(z_b(\mu), \mu) \nabla z_b(\mu) = 0.$$

Since $\{z_b(\mu) : \mu \in (0, 1]\}$ is bounded, the continuity of F and G implies that both $\|G(z, \mu)\|$ and $\|F(G(z, \mu))\|$ are bounded over T, say by m_1 and m_2, respectively. The transformation F is continuously differentiable, whence its Jacobian JF is uniformly bounded, say by m_3, on the compact set $cl(G(T))$.

Since $B = \sup\{\|\nabla_\mu G(z_b, \mu)\| : (z_b, \mu) \in T\} < \infty$, it follows that for all $\mu \in (0, \frac{1}{2}]$,

$$\begin{aligned}
\|\nabla_\mu H_b(z_b(\mu), \mu)\| &= \left\| \begin{array}{c} (1-\mu) JF(G(z_b(\mu), \mu))[\nabla_\mu G(z_b(\mu), \mu)] - F(G(z_b(\mu), \mu)) \\ -(1-\mu)\nabla_\mu G(z_b(\mu), \mu) + G(z_b(\mu), \mu) - b \end{array} \right\| \\
&\le \|(1-\mu) JF(G(z_b(\mu), \mu))[\nabla_\mu G(z_b(\mu), \mu)]\| + \|F(G(z_b(\mu), \mu))\| \\
&\quad + \|(1-\mu)\nabla_\mu G(z_b(\mu), \mu)\| + \|G(z_b(\mu), \mu)\| + \|b\| \\
&\le m_3 B + m_2 + B + m_1 + \|b\| \triangleq m_b,
\end{aligned}$$

which together with (4.6) and the compactness of $cl(T)$, implies

$$\|\nabla_\mu z_b(\mu)\| = \|-J_z H_b(z_b(\mu), \mu)^{-1} \nabla_\mu H_b(z_b(\mu), \mu)\| \le \frac{m_b}{\sigma}$$

for some $\sigma > 0$.

For any $\mu_1, \mu_2 \in (0, 1]$, it follows

$$\|z_b(\mu_1) - z_b(\mu_2)\| = \left\| \int_0^1 (\mu_1 - \mu_2) \nabla_\mu z_b(\mu_1 + t(\mu_2 - \mu_1)) dt \right\| \le \frac{m_b}{\sigma} |\mu_1 - \mu_2|,$$

which shows that $z_b(\mu)$ is convergent as $\mu \downarrow 0$.

Now, let z^* be a solution to the normal map equation; i.e., z^* satisfies $H_0(z, 0) = 0$. It is straightforward to deduce from the continuous differentiability of F and G that H_0 is locally Lipschitz at $(z^*, 0)$ over $\mathbb{E} \times \mathbb{R}_+$, say

$$\|H_0(z^* + x, \mu_1) - H_0(z^* + y, \mu_2)\| \le l_H(|\mu_1 - \mu_2| + \|x - y\|)$$

for all $(x, \mu_1), (y, \mu_2) \in \mathbb{E} \times \mathbb{R}_+$ with $\mu_1, \mu_2, \|x\|, \|y\| \le \bar{\varepsilon}$ and $\bar{\varepsilon} < 1$. Let $\bar{\sigma}$ be the multiplicative factor given in (4.6) with the compact set U as the closure of the union of all paths $\{z_b(\mu) : \mu \in (0, \bar{\varepsilon}]\}$ over $\{b : \|b\| \le l_H\}$.

Given an arbitrary $\varepsilon > 0$, let $\mu_\varepsilon = \min\{\bar{\varepsilon}, \frac{\varepsilon\bar{\sigma}}{l_H}\} > 0$. We have

$$\|H_0(z^*, \mu)\| = \|H_0(z^*, \mu) - H_0(z^*, 0)\| \le l_H \mu,$$

so that $b_\mu := \frac{1}{\mu} H_0(z^*, \mu)$ satisfies $\|b_\mu\| \le l_H$ for $\mu \le \mu_\varepsilon$. Note that $z^* = z_{b_\mu}(\mu)$. Differentiating $H_b(z_b(\mu), \mu) = 0$ with respect to b then gives

$$J_z H_b(z_b(\mu), \mu) J_b z_b(\mu) - \mu I = 0.$$

Thus, since $z_{tb_\mu}(\mu) \in U$ for all $t \in [0, 1]$, we may apply (4.6) to conclude that

$$\begin{aligned} \|z^* - z_0(\mu)\| &= \|z_{b_\mu}(\mu) - z_0(\mu)\| \\ &= \left\| \int_0^1 J_b z_{tb_\mu}(\mu) b_\mu dt \right\| \\ &= \left\| \int_0^1 \mu J_z H_b(z_{tb_\mu}(\mu), \mu)^{-1} b_\mu dt \right\| \\ &\le \frac{\mu}{\bar{\sigma}} \|b_\mu\| < \varepsilon \end{aligned}$$

for all $\mu \le \mu_\varepsilon$; hence $\|z^* - \lim_{\mu \downarrow 0} z_0(\mu)\| < \varepsilon$. Since $\varepsilon > 0$ is arbitrary, $z^* = \lim_{\mu \downarrow 0} z_0(\mu)$ is unique. □

Bibliography

[1] B. Chen and P. T. Harker. “A noninterior-point continuation method for linear complementarity problems”. In: *SIAM Journal on Matrix Analysis and Applications* 14 (1993), pp. 1168–1190.

[2] J. S. Chen and P. Tseng. “An unconstrained smooth minimization reformulation of the second-order cone complementarity problem”. In: *Mathematical Programming* 104 (2005), pp. 293–327.

[3] X. Chen and H. Qi. “Cartesian P-property and its applications to the semidefinite linear complementarity problem”. In: *Mathematical Programming* 106 (2006), pp. 177–201.

[4] X. Chen and P. Tseng. “Non-interior continuation methods for solving semidefinite complementarity problems”. In: *Mathematical Programming* 95 (2003), pp. 431–474.

[5] X. Chen and S. Xiang. “Computation of error bounds for P-matrix linear complementarity problem”. In: *Mathematical Programming* 106 (2006), pp. 513–525.

[6] X. Chen and S. Xiang. “Perturbation bound of P-matrix linear complementarity problems,” in: *SIAM Journal on Optimization* 18 (2007), pp. 1250–1265.

[7] C. B. Chua and Z. Li. “A barrier-based smoothing proximal point algorithm for NCPs over closed convex cones”. In: *SIAM Journal on Optimization* 23 (2012), pp. 745–769.

[8] C. B. Chua and P. Yi. “A continuation method for nonlinear complementarity problems over symmetric cones”. In: *SIAM Journal on Optimization* 20 (2010), pp. 2560–2583.

[9] F.H. Clarke. *Optimization and Nonsmooth Analysis*. New York: Wiley, 1983.

[10] B. C. Eaves and C. E. Lemke. “Equivalence of LCP and PLS”. In: *Mathematics of Operations Research* 6 (1981), pp. 475–484.

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

[12] J. Faraut and A. Korányi. *Analysis on symmetric cones*. New York: Oxford Press, 1994.

[13] M.C. Ferris and J. S. Pang. “Engineering and economic applications of complementarity probelms”. In: *SIAM Review* 39 (1997), pp. 669–713.

[14] M. Fiedler and V. Pták. "On matrices with non-positive off-diagonal elements and positive principal minors". In: *Czechoslovak Mathematical Journal* 12 (1962), pp. 382–400.

[15] A. Fischer. "A special Newton-type optimization method". In: *Optimization* 24 (1992), pp. 269–284.

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

[17] M. S. Gowda. "On the extended linear complementarity problem". In: *Mathematical Programming* 72 (1995), pp. 33–50.

[18] M. S. Gowda. "Reducing a monotone horizontal LCP to an LCP". In: *Applied Mathematics Letters* 8 (1995), pp. 97–100.

[19] M. S. Gowda and Y. Song. "On semidefinite linear complementarity problems". In: *Mathematical Programming* 88 (2000), pp. 575–585.

[20] O. Guler. "Generalized linear complementarity problem". In: *Mathematics of Operations Research* 20 (1995), pp. 441–448.

[21] S.-M. Guu. "Sign reversing and matrix classes". In: *Journal of Optimization Theory and Applications* 88 (1996), pp. 373–387.

[22] S.-M. Guu and R. W. Cottle. "On a subclass of P_0". In: *Linear Algebra and Its Applications* 223/224 (1995), pp. 325–335.

[23] P. T. Harker and J. S. Pang. "Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications". In: *Mathematical Programming* 48 (1990), pp. 161–220.

[24] P. Hartman and G. Stampacchia. "On some nonlinear elliptic differential functional equations". In: *Acta Mathematica* 115 (1966), pp. 271–310.

[25] Z. Huang and J. Han. "Non-interior continuation method for solving the monotone semidefinite complementarity problem". In: *Applied Mathematics and Optimization* 47 (2003), pp. 195–211.

[26] C. Kanzow. "An unconstrained optimization technique for large-scale linearly constrained convex minimization problem". In: *Computing* 53 (1994), pp. 101–117.

[27] C. Kanzow. "Some noninterior continuation methods for linear complementarity problems". In: *SIAM Journal on Matrix Analysis and Applications* 17 (1996), pp. 851–868.

[28] L. Kong, L. Tunçel and N. Xiu. "Monotonicity of Lowner operators and its applications to symmetric cone complementarity problems". In: *Mathematical Programming* 133 (2012), pp. 327–336.

[29] G. E. Lemke and J. T. Howson. "Equilibrium points of bimatrix games". In: *SIAM Journal on Applied Mathematics* 12 (1964), pp. 413–423.

[30] X. Liu and F. A. Potra. “Corrector-predictor methods for sufficient linear complementarity problems in a wide neighborhood of the central path”. In: *SIAM Journal on Optimization* 17 (2006), pp. 871–890.

[31] A. Lohne. “A characterization of maximal monotone operators”. In: *Set-Valued Analysis* 16 (2008), pp. 693–700.

[32] M. Fukushima, Z. Q. Luo and P. Tseng. “Smoothing functions for second-order-cone complementarity problems”. In: *SIAM Journal on Optimization* 12 (2001), pp. 436–460.

[33] Aganagic M. “Contributions to Complementarity Theory”. PhD Thesis. Stanford University, 1978.

[34] O. L. Manasarian and J. S. Pang. “The extended linear complementarity problem”. In: *SIAM Journal on Matrix Analysis and Applications* 16 (1995), pp. 359–368.

[35] O. L. Mangasarian. *Nonlinear Programming*. New York: McGraw-Hill, 1969.

[36] A. Friedlander, J. M. Martínez and S. A. Santos. “On the resolution of linearly constrained convex minimization problems”. In: *SIAM Journal on Optimization* 4 (1994), pp. 331–339.

[37] A. Friedlander, J. M. Martínez and S. A. Santos. “Solution of linear complementarity problems using minimization with simple bounds”. In: *Journal of Global Optimization* 6 (1995), pp. 253–267.

[38] G. J. Minty. “Monotone (nonlinear) operators in Hilbert space”. In: *Duke Mathematical Journal* 29 (1962), pp. 314–346.

[39] R. D. C. Monteiro and J. S. Pang. “Properties on an interior-point mapping for mixed complementarity problems”. In: *Mathematics of Operations Research* 21 (1996), pp. 629–654.

[40] J. J. Moreau. “Décomposition orthogonale d’un espace hilbertien selon deux cones mutuellement polaires”. In: *Comptes Rendus de l’Académie des Sciences* 255 (1962), pp. 238–240.

[41] K. Murty. *Linear Complementarity, Linear and Non-linear Programming*. Internet Edition, 1995.

[42] M. Kojima, N. Megiddo, T. Noma and A. Yoshise. *Lecture Notes in Computer Science: A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems*. New York: Springer, 1991.

[43] R. W. Cottle, J. S. Pang and R. E. Stone. *The Linear Complementarity Problem*. Boston: Academic Press, Inc., 1992.

[44] F. A. Potra. *Sufficient weighted complementarity problems*. Technical Report. University of Maryland Baltimore County, 2014.

[45] F. A. Potra. "Weighted Complementarity Problems". In: *SIAM Journal on Optimization* 22 (2012), pp. 1634–1654.

[46] X. Chen, L. Qi and P. Tseng. "Analysis of nonsmooth symmetric-matrix functions with applications to semidefinite complementarity problems". In: *SIAM Journal on Optimization* 13 (2003), pp. 960–985.

[47] R. T. Rockafellar. "Monotone operators and the proximal point algorithm". In: *SIAM Journal on Control and Optimization* 14 (1976), pp. 877–898.

[48] R. T. Rockafellar. "On the maxiaml monotonicity of subdifferential mappings". In: *Pacific Journal of Mathematics* 33 (1970), pp. 209–216.

[49] M. Shida and S. Shindoh. *Monotone semidefinite complementarity problems*. Technical Report. Tokyo Institute of Technology, Tokyo, Japan, 1996.

[50] S. Smale. *Algorithms for solving equations*. Proceedings of the International Congress of Mathematics. Berkeley, CA, 1986, pp. 172–195.

[51] M. V. Solodov. "Some optimization reformulation of the extended linear complementarity problem". In: *Computational Optimization and Applications* 13 (1999), pp. 187–200.

[52] D. Sun and L. Qi. "On NCP-functions". In: *Computational Optimization and Applications* 13 (1999), pp. 201–220.

[53] D. Sun and L. Qi. "Solving variational inequality problems via smoothing-nonsmooth reformulations". In: *Journal of Computational and Applied Mathematics* 129 (2001), pp. 37–62.

[54] L. Kong, J. Sun and N. Xiu. "A regularized smoothing Newton method for symmetric cone complementarity problems". In: *SIAM Journal on Optimization* 19 (2008), pp. 1028–1047.

[55] L. Qi, D. Sun and J. Sun. "A squared smoothing Newton method for nonsmooth matrix equations and its applications in semidefinite optimization problems". In: *SIAM Journal on Optimization* 14 (2003), pp. 783–806.

[56] X. Chen, D. Sun and J. Sun. "Complementarity functions and numerical experiments on some smoothing Newton methods for second-order-cone complementarity problems". In: *Computational Optimization and Applications* 25 (2003), pp. 39–56.

[57] M. S. Gowda, R. Sznajder and J. Tao. "Some P-properties for linear transformations on Euclidean Jordan algebras". In: *Linear Algebra and Its Applications* 393 (2004), pp. 203–232.

[58] R. Sznajder and M. S. Gowda. "Generalizations of P_0 and P-properties; extended vertical and horizontal linear complementarity problems". In: *Linear Algebra and Its Applications* 223-224 (1995), pp. 695–715.

[59] J. Tao and M. S. Gowda. "Some P-properties for nonlinear transformations on Euclidean Jordan algebras". In: *Mathematics of Operations Research* 30 (2005), pp. 985–1004.

[60] H. Samelson, R. M. Thrall and O. Wesler. "A partition theorem for Euclidean space". In: *Proceedings of American Mathematical Society* 9 (1958), pp. 805–807.

[61] H. Väliaho. "P_*-matrices are just sufficient". In: *Linear Algebra and Its Applications* 239 (1996), pp. 103–108.

[62] L. Walras. *Elements of Pure Economies.* London: Allen and Unwin, 1954.

[63] J. G. Wardrop. "Some theoretical aspects of road traffic research". In: *Proceeding of the Institute of Civil Engineers, Part II* (1952), pp. 325–378.

[64] S. Hayashi, N. Yamashita and M. Fukushima. "A combined smoothing and regularization method for monotone second-order cone complementarity problems". In: *SIAM Journal on Optimization* 15 (2005), pp. 593–615.

[65] Y. Ye. "A fully polynomial-time approximation algorithm for computing a stationary point of the general complementarity problem". In: *Mathematics of Operations Research* 18 (1994), pp. 334–345.

[66] A. Yoshise. "Interior point trajectories and a homogeneous model for nonlinear complementarity problems over symmetric cones". In: *SIAM Journal on Optimization* 17 (2006), pp. 1129–1153.

[67] Y. Zhang. "On the convergence of a class of infeasible interior-point algorithms for the horizontal linear complementarity problem". In: *SIAM Journal on Optimization* 4 (1994), pp. 208–227.