CENTER FOR
PARALLEL OPTIMIZATION

COMPUTABLE ERROR BOUNDS IN MATHEMATICAL PROGRAMMING

by

Jun Ren

Computer Sciences Technical Report #1173

August 1993

COMPUTABLE ERROR BOUNDS IN MATHEMATICAL PROGRAMMING

By

Jun Ren

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy
(Computer Sciences)

at the

UNIVERSITY OF WISCONSIN – MADISON

1993

Abstract

This thesis establishes error bounds for linear complementarity problems, quadratic programs, strongly monotone nonlinear complementarity problems, and strongly convex programs. After obtaining these theoretical error bounds, we apply them to approximate solutions generated by various algorithms for solving these problems. This provides bounds on how far each approximate solution is from the set of exact solutions of each problem.

After surveying the relevant literature, we prove a new global error bound for monotone linear complementarity problems. We also extend another global error bound for the linear complementarity problem to cover the cases where the underlying matrix is an R_0-matrix and give a new simple proof for this extension. We compare the new error bounds with existing ones and obtain relations between them. We establish both an upper and a lower bound locally on the distance to the solution set of any linear complementarity problem from an arbitrary point. We also consider other problems such as nonlinear complementarity problems and convex programming problems. For the strongly monotone nonlinear complementarity problem, we obtain that a natural residual is a global error bound. In addition an equivalence relation has been established between a natural residual and the implicit Lagrangian function for a general nonlinear complementarity problem. Immediate consequences are that the square root of the implicit

Lagrangian function is a local error bound for any linear complementarity problem, and is a global error bound for strongly monotone nonlinear complementarity problems.

We further consider various computational algorithms such as exterior penalty, interior penalty, proximal point, and augmented Lagrangian methods for solving linear programs, convex quadratic programs, strongly monotone complementarity problems and strongly convex programs. We bound the distance between any approximate solution obtained by any one of the methods and the set of exact solutions for the problem being solved. For many cases, we prove that this distance depends only on computable quantities such as the size of the penalty parameters, the amount of the constraint violation or the amount of violation of the complementarity condition. This makes it possible for these error bounds to be used as termination criteria for these algorithms.

Acknowledgements

I wish to express my deepest gratitude to my advisor Professor Olvi Mangasarian for his advice, help , support and encouragement. Without his guidance this thesis could not have been written.

I wish to express my appreciation to Professor Stephen M. Robinson who encouraged me to go for the Ph.D. He was instrumental in my coming to Madison and has always been generous with his time, advice and help. I am thankful to Professors Robert R. Meyer and Michael C. Ferris for reading a draft of this thesis and Professor Renato De Leone for being a member of the oral exam committee. I also would like to thank Professor Liqun Qi for arranging my coming to Madison.

My debt to my wife Jinwen Huang for support throughout my doctoral studies is enormous. She has constantly encouraged me for my Ph.D. studies and has sacrificed her time and energy for my work. I can only hope to match it with my gratitude. I also would like to thank my father Guojun Ren and mother Yaoming Chen who, even from my early school days, have always encouraged and supported me in my studies.

This research was partially funded by National Science Foundation Grant CCR-9101801 and Air Force Office of Scientific Research Grant AFOSR-89-0410.

Contents

Chapter 1

Overview and Summary

1.1 Introduction

Error bounds are useful tools in establishing the accuracy of approximate solutions to mathematical programs, inequalities and complementarity problem. New error bounds and their mathematical theory have led to new algorithms and establishment of the convergence and linear convergence rate for these algorithms. For example, Pang [Pa86] obtained convergence results for the inexact Newton method for the nonlinear complementarity problem. Recently, by generating a new local error bound, Luo and Tseng [LuT92a] have established global convergence results with locally linear rate for matrix splitting algorithms for the affine variational inequality problem. By obtaining another new global error bound, Mangasarian [Ma91] completed a global linear convergence proof for the same problem by using a new global error bound. In addition, there are a number of new other convergence results for other algorithms based on error bounds ([LuT91] [LuT92b] [LuT92c] [LuT92d]).

Error bounds for linear inequalities were first introduced by Hoffman [Hof52].

For a given solvable linear inequality system, Hoffman proved that for an arbitrary point, the distance between this point and the solution set of the inequality system can be bounded by a measure of the violation of the inequalities at the point. Different error bounds were subsequently obtained for linear programming ([Rob73], [MaS87]). Such error bounds can serve as termination criteria for iterative algorithms and can be used to estimate the amount of error allowable in an inexact computation of the iterates that converge to a solution of a mathematical program. Recently more error bounds for various optimization problems have been established ([Rob81] [Pa86] [MaS86] [MaPa90] [FeM91] [Gül91] [GHR92] [Li91] [LuT92a] [LuT92b] [MaD88] [Ma90] [Ma91] [LMRS92]).

New error bounds have been found to be very important and useful in proving the convergence and linear convergence rate for various algorithms. Also they are a source and a tool for generating new algorithms.

1.2 Error Bounds and Algorithm-Generated Error Bounds

In this thesis, we develop new error bounds for mathematical programming problems and analyze relationships among different error bounds and error residuals. We then apply these error bounds to various algorithms to obtain a closeness estimate between an inexact solution obtained by an algorithm and the true solution set of the problem we want to solve. We prove that for many algorithms these estimates can be expressed in terms of computable quantities such as the size of the penalty parameters and the violation of the constraints. Therefore, these estimates can serve as termination criteria for iterative algorithms.

In Chapter 2, we develop new error bounds for the linear complementarity problem. Based on a local error bound [LuT92a] and a global error bound [MaS86], a new improved global error bound is established by properly combining these two bounds. For the linear complementarity problem with an R_0 coefficient matrix, we show that a natural residual is a global error bound. The novel idea of the proof consists in assuming that there are points that cannot be bounded by this residual and then proving that this leads to a contradiction. This turns out to be a very simple proof. This result extends the results in [MaPa90] [LMRS92]. We also show that Mangasarian and Shiau error bound holds [MaS86] for an LCP with a R_0 matrix besides a monotone LCP. For the general (indefinite) linear complementarity problem, we show that any known error residual fails to be an error bound without multiplying it by a term involving the norm of the point under consideration. In addition, we investigate relations among different error bounds and error residuals. Relations between the new error bounds we developed and exiting error bounds are given.

In Chapter 3, a new global error bound for the strongly monotone nonlinear complementarity problem is obtained. This result is a natural extension of a well-known natural error bound for the strongly monotone linear complementarity problem [Pa86]. We also generalize the Mangasarian and Shiau error bound for the linear complementarity problem to cover the strongly monotone nonlinear complementarity problem.

In Chapter 4, we discuss the implicit Lagrangian function [MaS92] as an error bound. We prove that the square root of this function can be bounded both from above and below by a natural error residual. This shows that the square root of the implicit Lagrangian is equivalent to the natural residual. Consequently, it is a local error bound for any linear complementarity problem, and becomes a global error bound for the strongly monotone nonlinear complementarity problem.

In Chapter 5, we use the error bounds that we have developed, to measure the closeness between each inexact solution that is generated by certain commonly-used algorithms and the true solution set of the problem we want to solve. We apply exterior penalty, interior penalty, augmented Lagrangian and proximal point methods to problems such as linear programs, convex quadratic programs, strongly monotone nonlinear complementarity problems and strongly convex programs. By using the optimality conditions satisfied by inexact solutions generated by these approaches, we are able to eliminate the dual variables from the error bound. Thus we obtain error bounds involving only the primal variables at an inexact solution. In many cases, these actual bounds depend only on the size of certain parameters such as penalty parameters or violations of the constraints or violations of the complementarity condition. These error bounds can be used as termination criteria for these algorithms because they are computable and bound the distance between the approximate solution and the real solution set of the original problem.

In Chapter 6, we end with a brief summary, some open questions and some directions for new research.

1.3 Notation

All scalars, vectors and matrices in this thesis are real. We define below the notation used throughout this thesis.

1. $I\!R^n$ denotes the n-dimensional real space, and $I\!R^n_+$ denotes the nonnegative orthant, or the set of all points in $I\!R^n$ with nonnegative components.

2. $x \in I\!R^n$ denotes an n-dimensional column or row vector. For vectors $x \in I\!R^n$ and $y \in I\!R^m$, (x, y) denotes a vector in $I\!R^{n+m}$ with the i-th component

$(x, y)_i = x_i$ if $i \leq n$; $(x, y)_i = y_{i-n}$ otherwise.

3. A subscript is used to denote a component of a vector, e.g. x_i is the i-th component of a vector $x \in I\!R^n$.

4. For a vector $x \in I\!R^n$, x_+ is a n-dimensional vector with $(x_+)_i = \max\{x_i, 0\}$, $i = 1, \ldots, n$, and $|x|$ with $|x|_i = |x_i|$, $i = 1, \ldots, n$. We say $x > 0$ for a vector $x \in I\!R^n$ if $x_i > 0, \forall i = 1, \ldots, n$. Similarly, $x \geq 0$ if $x_i \geq 0, \forall i = 1, \ldots, n$.

5. For two vectors $x \in I\!R^n$ and $y \in I\!R^n$, $\min\{x,\ y\}$ is an vector in $I\!R^n$ with $(\min\{x,\ y\})_i = \min\{x_i,\ y_i\}$, $i = 1, \ldots, n$.

6. Superscripts are used to differentiate between vectors, e.g. x^k and x^{k+1}.

7. The scalar product of two vectors x and y in $I\!R^n$ is denoted by xy.

8. The Euclidean or 2-norm of an $x \in I\!R^n, (xx)^{1/2}$, is denoted by $\|\cdot\|_2$, while the 1-norm, $\sum_{i=1}^n |x_i|$ by $\|\cdot\|_1$, and an arbitrary norm by $\|\cdot\|$.

9. $I\!R^{m \times n}$ denotes the set of all real $m \times n$ matrices. All matrices are denoted by upper case letters.

10. For a matrix $A \in I\!R^{m \times n}$, A^T denotes the transpose of $A \in I\!R^{n \times m}$.

11. A matrix $A \in I\!R^{n \times n}$ is positive definite if $xAx > 0, \forall x \neq 0 \in I\!R^n$, positive semi-definite if $xAx \geq 0, \forall x \in I\!R^n$, a P-matrix if all its principal minors are positive and a nondegenerate matrix if they are nonzero.

12. For a function $F(x)$ from $I\!R^n$ to $I\!R^m$, $\nabla F(x)$ denotes its $m \times n$ Jacobian.

13. A matrix $A \in I\!R^{n \times n}$ is said to be in class R_0 if $Mx \geq 0, x \geq 0, xMx = 0$ implies that $x = 0$.

14. A norm $\|\cdot\|$ is called a monotonic if $\|x\| = \|\,|x|\,\|$

Chapter 2

New Global Error Bounds for the Linear Complementarity Problem (LCP)

2.1 Introduction

In this chapter, we develop various error bounds for the linear complementarity problem. We also establish a number of relationships among different error residuals that are used to bound the distance between an arbitrary point and the solution set of the problem.

We consider the classical linear complementarity problem ([Mur88] [CoPS92]) of finding an x in the n-dimensional real space $I\!R^n$ such that

$$Mx + q \geq 0, \quad x \geq 0, \quad x(Mx + q) = 0, \tag{2.1.1}$$

where $M \in I\!R^{n\times n}$ and $q \in I\!R^n$. We denote this problem by LCP(M, q) for short.

Let the solution set be

$$\bar{X} := \{x \mid Mx + q \geq 0,\ x \geq 0,\ x(Mx+q) = 0\}. \tag{2.1.2}$$

We assume that $\bar{X}$ is nonempty. Define the *natural* residual ([Pa86], [LuT92a], [Ma91]

$$r(x) := \|x - (x - Mx - q)_+\|. \tag{2.1.3}$$

We also introduce another residual ([MaS86], [Ma90])

$$s(x) := \|(-Mx - q,\ -x,\ x(Mx+q))_+\|. \tag{2.1.4}$$

Note that $x \in \bar{X}$ is equivalent to $r(x) = 0$ or $s(x) = 0$. We first give the following definitions of a residual and error bound for the LCP.

Definition 2.1.1 *Let* $e : I\!R^n \to I\!R$.

1. $e(x)$ *is a* ***residual*** *for the LCP(M, q) if* $e(x) \geq 0$, *for all* $x \in I\!R^n$, *and* $e(x) = 0$ *if and only if* x *solves LCP(M, q).*

2. $e(x)$ *is a global (local)* ***error bound*** *for the LCP(M, q) if it is a residual such that there exists some constant* τ *(and* $\epsilon > 0$*) such that for each* $x \in I\!R^n$ *(when* $e(x) \leq \epsilon$*)*

   $$\|x - \bar{x}(x)\| \leq \tau e(x)$$

   where $\bar{x}(x)$ *is a closest solution of LCP(M, q) to* x *under the norm* $\|\cdot\|$.

It is obvious that if $e(x)$ is a global error bound, then it is a local error bound. In addition Definition (2.1.1) can be extended to cases where the problem could be any mathematical programming problem such as a linear program, a nonlinear program or a nonlinear complementarity problem.

2.2 The Residual $r(x)$ as an Error Bound

There have been a number of recent papers ([LMRS92], [LuT92a], [Ma90], [Ma91], [MaPa90], [MaS86], [Pa86]) that use $r(x)$ and $s(x)$ to bound the distance from a point $x \in I\!R^n$ to the solution set $\bar{X}$. It turns out that some properties of M such as positive definiteness and positive semi-definiteness play important roles in generating error bounds. For an indefinite matrix M, we still do not have a global error bound that does not involve the norm of the point in question ([Ma91]). We begin by giving some basic error bounds for the LCP.

The first global error bound using $r(x)$ was introduced by Pang ([Pa86]) for LCP(M, q) with a positive definite M. We state it in the following lemma.

Lemma 2.2.1 *[Pa86] Let $M \in I\!R^{n\times n}$ be positive definite, then there exists a constant $\lambda > 0$ such that*

$$\|x - \bar{x}\| \leq \lambda r(x), \qquad \forall x \in R^n \tag{2.2.1}$$

where $\bar{x}$ is the unique solution of the LCP. If $\|\cdot\|$ denotes the 2-norm in (2.2.1) and in the definition (2.1.3) of $r(x)$, then

$$\lambda = 1 + \frac{\|I - M\|}{\mu},$$

where μ is the smallest eigenvalue of the matrix $\frac{M+M^T}{2}$.

Proof. See [Pa86]. **Q.E.D.**

Remark. Lemma 2.2.1 also provides a computable λ in term of the matrix M which is not the case in general. Hence it is possible to obtain an explicit upper bound on the distance between any point and a closest solution of a positive definite LCP. This can be used to decide how good an approximate solution is based on this error. However, Lemma 2.2.1 is too restrictive because of the positive definiteness assumption on M and it fails to hold for a positive semi-definite M ([MaS86]).

Although the residual $r(x)$ cannot provide a global error bound for an indefinite LCP(M, q), it does provide a local error bound for it ([Rob81], [LuT92a]). This result We will use this result later on. So we state it in the following lemma.

Lemma 2.2.2 *[Rob81] [LuT92a] Let $M \in I\!R^{n\times n}$, then there exist $\epsilon > 0$ and $\tau > 0$ such that*

$$\|x - \bar{x}(x)\| \leq \tau r(x), \quad \forall x,\ r(x) \leq \epsilon \tag{2.2.2}$$

where $\bar{x}(x)$ is a closest solution of LCP(M, q) to x under the norm $\|\cdot\|$.

Proof. See [LuT92a] and [Rob81] **Q.E.D.**

It should be pointed that by using Lemma 2.2.2, Luo and Tseng ([LuT91] [LuT92a] [LuT92b] [LuT92c], [LuT92d]) recently established the convergence and locally linear convergence rate results for a number of algorithms for various mathematical problems. In addition, by using a different new global error bound, Mangasarian ([Ma91]) globalized the locally linear convergence for the matrix splitting algorithm for affine variational inequality problems. The key to achieving all these results is the availability of a measure of the error reduction ratio between consecutive iterates obtained by using an error bound. As a result, convergence and convergence rate results can be obtained for various algorithms. New error bounds are an important tool and source for such results.

Recently, it was found that $r(x)$ provides a global error bound for larger classes of matrices M and not only for positive definite matrices. For example, $r(x)$ is a global error bound for P-matrices M ([MaPa90]) and more generally, for nondegenerate matrices M ([LMRS92]). In this work, we further extend these results to an even larger class of matrices M. Specifically, we prove that $r(x)$ is a global error bound for R_0-matrices M. This class contains all these three other classes ([CoPS92]). We give a completely different and simpler proof by using Lemma 2.2.2. First we define the class of R_0-matrices ([CoPS92]).

Definition 2.2.1 *An $M \in \mathbb{R}^{n\times n}$ is called an R_0-matrix if the LCP(M, 0) has zero as its unique solution.*

Theorem 2.2.1 *Let $M \in \mathbb{R}^{n\times n}$ be an R_0-matrix. Then there exists $\tau > 0$ such that for each $x \in \mathbb{R}^{n\times n}$*

$$\|x - \bar{x}(x)\| \leq \tau r(x), \tag{2.2.3}$$

where $\bar{x}(x)$ is a closest solution of LCP(M, q) to x under the norm $\|\cdot\|$.

Proof. Assume that the theorem is false. Then for each integer k, there exists an x^k such that (2.2.3) is violated, i.e.

$$\|x^k - \bar{x}(x^k)\| > kr(x^k),$$

where $\bar{x}(x^k)$ is a closest solution of LCP(M, q) to x^k under the norm $\|\cdot\|$. In particular choose now a fixed solution $\bar{x}$ such that

$$\|x^k - \bar{x}\| \geq \|x^k - \bar{x}(x^k)\| > kr(x^k). \tag{2.2.4}$$

Since $r(x)$ is a local error bound, it follows by Lemma 2.2.2, there exist $K > 0$ and $\epsilon > 0$ such that $r(x^k) > \epsilon$, for $k > K$. Otherwise, we would have for all $K > 0, \epsilon > 0$, there exists some $k > K$ such that $r(x^k) \leq \epsilon$. This implies, because $r(x)$ is a local error bound, that

$$\frac{\tau}{k}\|x - \bar{x}(x^k)\| > \tau r(x^k) \geq \|x - \bar{x}(x^k)\|$$

where the first inequality follows from (2.2.4). This leads to the contradiction $\frac{\tau}{k} \geq 1$ as $k \to \infty$, where τ is defined in (2.2.2). Hence the right hand side of (2.2.4) goes to infinity as k goes to infinity and so does the left hand side since it is bigger. Therefore, $\|x^k\|$ goes to infinity. Without loss of generality, let

$$\lim_{k\to\infty} \frac{x^k}{\|x^k\|} = s$$

Note that $\|s\| = 1$. Divide both sides of (2.2.4) by $\|x^k\|$ and let k goes to infinity to obtain

$$\begin{aligned}
1 &= \lim_{k\to\infty} \frac{\|x^k - \bar{x}\|}{\|x^k\|} \\
&\geq \lim_{k\to\infty} k * \frac{r(x^k)}{\|x^k\|} \\
&= \lim_{k\to\infty} k * \|\min\{\frac{x^k}{\|x^k\|}, \frac{Mx^k + q}{\|x^k\|}\}\| \\
&= \lim_{k\to\infty} k * \|\min\{s, Ms\}\|.
\end{aligned}$$

Therefore $\min\{s, Ms\} = 0$. This is equivalent to the LCP(M, 0) having a nonzero solution s. This contradicts the assumption that $M \in R_0$. **Q.E.D.**

Previous results of [Pa86], [MaPa90], [LMRS92] follow as a corollary of Theorem 2.2.1.

Corollary 2.2.1 *Let $M \in I\!R^{n\times n}$ be nondegenerate. Then there exists a $\tau > 0$ such that for any $x \in I\!R^n$*

$$\|x - \bar{x}(x)\| \leq \tau r(x),$$

where $\bar{x}(x)$ is a closest solution of LCP(M, q) to x under the norm $\|\cdot\|$.

Proof. Since each nondegenerate matrix is a R_0-matrix ([CoPS92]), the proof follows from Theorem 2.2.1. **Q.E.D.**

2.3 Variations of $r(x)$ and $s(x)$ as Global Error Bounds

In this section, we discuss variations of $s(x)$ and $r(x)$ as global bounds on the distance between points $x \in I\!R^n$ and the solution set of LCP(M, q) for wider classes of matrices M. It stated in the previous section that $r(x)$ by itself fails to provide a global error bound for LCP(M, q) with a positive semi-definite M. Therefore, other types of residuals are required in order to obtain a global error bound for LCP(M, q) with positive semi-definite M. The residual $s(x)$ is one such residual. Mangasarian and Shiau ([MaS86]) established the first global error bound for LCP(M, q) with a positive semi-definite M. This result is summarized in the following lemma which will be used throughout this thesis.

Lemma 2.3.1 *[MaS86] [Ma90] Let $M \in I\!R^{n\times n}$ be positive semi-definite, then there exists a constant $\gamma > 0$ such that*

$$\|x - \bar{x}(x)\| \leq \gamma(s(x) + s(x)^{\frac{1}{2}}), \tag{2.3.1}$$

where $\bar{x}(x)$ is a closest solution of LCP(M, q) to x under the norm $\|\cdot\|$. In particular, if the LCP (M, q) has a nondegenerate solution, that is $\bar{x}+M\bar{x}+q > 0$ for some solution $\bar{x}$, then we have the simpler error bound

$$\|x - \bar{x}(x)\| \leq \gamma s(x)$$

Proof. See [MaS86] and [Ma90]. **Q.E.D.**

Remark. Lemma 2.3.1 fails to hold if we delete the term $s(x)^{1/2}$ in (2.3.1). See Example 2.9 in [MaS86]. The residual $s(x)$ alone provides a global error bound only if M is positive semi-definite and LCP(M, q) has a nondegenerate solution.

On the other hand, $s(x)^{1/2}$ does not provide an error bound by itself as can be seen from the following example. Let

$$M = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \qquad q = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

It is easy to see that $\bar{x} = (0,\ 0)$ is the unique solution for this LCP where M is positive semi-definite. Now let sequence $x^k = (k,\ 0),\ k = 1, 2, \ldots.$ Then $Mx^k + q = (1,\ 0)$, and under the 1-norm

$$s(x^k) = \|(-Mx^k - q,\ -x^k,\ x^k(Mx^k + q))_+\| = k.$$

So

$$\lim_{k\to\infty} \frac{\|x - \bar{x}\|}{s(x)^{1/2}} = \lim_{k\to\infty} \frac{k}{\sqrt{k}} = +\infty.$$

The following theorem shows that even though $r(x)$ and $s(x)$ individually fail to be an error bound for the positive semi-definite case, their sum does indeed provide a new global error bound. Moreover, the following theorem says that $r(x) + s(x)$ covers no fewer classes of matrices M than $s(x) + s(x)^{\frac{1}{2}}$ does as a global error bound for LCP(M, q), even though $s(x) + s(x)^{\frac{1}{2}} \not\leq c(r(x) + s(x))$ for some constant $c > 0$ and all $x \in I\!R^n$.

Theorem 2.3.1 *Let the residual $s(x) + s(x)^{\frac{1}{2}}$ be a global error bound for LCP(M, q) for some M, that is, there exists a τ such that*

$$\|x - \bar{x}(x)\| \leq \tau(s(x) + s(x)^{\frac{1}{2}}),$$

where $\bar{x}(x)$ is a closest solution of LCP(M, q) to x under the norm $\|\cdot\|$. Then there exists a constant $\bar{\tau} > 0$ such that

$$\|x - \bar{x}(x)\| \leq \bar{\tau}(r(x) + s(x)). \tag{2.3.2}$$

Proof. By Lemma 2.2.2, there exist $\epsilon > 0$ and $\tau_1 > 0$ such that if $r(x) \leq r(x) + s(x) \leq \epsilon$, then

$$\|x - \bar{x}(x)\| \leq \tau_1 r(x) \leq \tau_1 (r(x) + s(x)).$$

Else if $r(x) + s(x) \geq \epsilon$, consider the following two cases:

Case 1 $r(x) \geq s(x)$. In this case $r(x) \geq \epsilon/2$, and it is easy to see $r(x) \geq [(\epsilon/2)s(x)]^{1/2}$ together with $s(x) + s(x)^{\frac{1}{2}}$ being a global error bound. Then it follows that

$$\begin{aligned} \|x - \bar{x}(x)\| &\leq \tau(s(x) + s(x)^{\frac{1}{2}}) \\ &\leq \tau(s(x) + (2/\epsilon)^{1/2} r(x)) \\ &\leq \tau max\{1, (2/\epsilon)^{1/2}\}(s(x) + r(x)) \\ &\leq \bar{\tau}(s(x) + r(x)), \end{aligned}$$

where $\bar{x}(x)$ is defined as above and $\bar{\tau} = \max\{\tau_1, \tau(1 + (2/\epsilon)^{1/2})\}$.

Case 2 $r(x) \leq s(x)$, it follows that $s(x) \geq \epsilon/2$, so $s(x) \geq [(\epsilon/2)s(x)]^{1/2}$, hence,

$$\begin{aligned} \|x - \bar{x}(x)\| &\leq \tau(s(x) + s(x)^{\frac{1}{2}}) \\ &\leq \tau(s(x) + (2/\epsilon)^{1/2} s(x)) \\ &\leq \bar{\tau} s(x) \\ &\leq \bar{\tau}(r(x) + s(x)). \end{aligned}$$

Therefore (2.3.2) holds for all $x \in I\!R^n$. **Q.E.D.**

Corollary 2.3.1 *Let $M \in I\!R^{n \times n}$ be positive semi-definite. Then $r(x) + s(x)$ is a global error bound for LCP(M, q).*

Proof. By Lemma 2.3.1 and Theorem 2.3.1, the corollary follows immediately.

In fact, there are cases of LCP where $r(x) + s(x)$ is a global error bound, but $s(x) + s(x)^{\frac{1}{2}}$ is not.

Theorem 2.3.2 *The residual $r(x)+s(x)$ is a global error bound for a wider class of LCPs than the residual $s(x)+s(x)^{\frac{1}{2}}$.*

Proof. By Theorem 2.3.1, it is sufficient to prove that there is an LCP(M, q) for which $r(x)+s(x)$ is a global error bound, but $s(x)+s(x)^{\frac{1}{2}}$ is not. Let

$$M = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad q = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The solution set $\bar{X}$ of LCP(M, q) is $\{x \mid x_1 = x_3 = 0, x_2 \geq 0, \text{ or } x_2 = x_3 = 0, x_1 \geq 0\}$. First we prove that $r(x)+s(x)$ is a global error bound for this LCP. By computing the $r(x)$ and $\|x - \bar{x}(x)\|$, it follows that

$$\begin{aligned} r^2(x) &= \|\min\{x,\ Mx\}\|_2^2 = (\min\{x_1,\ x_2\})^2 + (-x_2)_+^2 + x_3^2, \\ \|x - \bar{x}(x)\|_2^2 &= \min\{x_1^2 + (-x_2)_+^2 + x_3^2,\ (-x_1)_+^2 + x_2^2 + x_3^2\}. \end{aligned} \tag{2.3.3}$$

First consider $x_1 \leq x_2$. Obviously

$$r^2(x) = x_1^2 + (-x_2)_+^2 + x_3^2.$$

Furthermore,

$$\begin{aligned} &\text{If } \ x_2 \geq x_1 \geq 0: \ \|x - \bar{x}(x)\|_2^2 = \min\{x_1^2 + x_3^2, x_2^2 + x_3^2\} \leq x_1^2 + x_3^2 = r(x)^2; \\ &\text{If } \ x_2 \geq 0 \geq x_1: \ \|x - \bar{x}(x)\|_2^2 = \min\{x_1^2 + x_3^2, x_1^2 + x_2^2 + x_3^2\} \leq x_1^2 + x_3^2 \leq r(x)^2; \\ &\text{If } \ 0 \geq x_2 \geq x_1: \ \|x - \bar{x}(x)\|_2^2 = \min\{x_1^2 + x_2^2 + x_3^2, x_1^2 + x_2^2 + x_3^2\} = r(x)^2. \end{aligned}$$

Hence for the these cases above

$$\|x - \bar{x}(x)\|^2 \leq r^2(x).$$

For the case where $x_1 \geq x_2$, (2.3.3) gives

$$r^2(x) = x_2^2 + (-x_2)_+^2 + x_3^2$$

and

$$\begin{aligned}
&\text{If} \quad x_1 \geq x_2 \geq 0: \quad \|x - \bar{x}(x)\|_2^2 = \min\{x_1^2 + x_3^2, x_2^2 + x_3^2\} \leq x_2^2 + x_3^2 = r(x)^2; \\
&\text{If} \quad x_1 \geq 0 \geq x_2: \quad \|x - \bar{x}(x)\|_2^2 = \min\{x_1^2 + x_2^2 + x_3^2, x_2^2 + x_3^2\} \leq x_2^2 + x_3^2 \leq r(x)^2; \\
&\text{If} \quad 0 \geq x_1 \geq x_2: \quad \|x - \bar{x}(x)\|_2^2 = \min\{x_1^2 + x_2^2 + x_3^2, x_1^2 + x_2^2 + x_3^2\} \leq r(x)^2.
\end{aligned}$$

Hence $r(x) + s(x)$ is a global error bound for this LCP. On the other hand, take $x^k = (-k^{-4}, k^2, k^{-1})$, then $Mx^k + q = (k^2, 0, k^{-1})$ and it follows that

$$\begin{aligned}
s(x^k) &= \|(-Mx^k - q, -x^k, x^k(Mx^k + q))_+\|_2 = k^{-4} \\
\|x^k - \bar{x}(x^k)\| &= \|k^{-4}, k^{-1}\|_2 > k^{-1}.
\end{aligned}$$

Hence $s(x) + s(x)^{\frac{1}{2}}$ is not an error bound for this LCP. **Q.E.D.**

Remark. Both error bounds (2.3.1) and (2.3.2) fail to hold for indefinite symmetric M as can be seen from the following example.

Let

$$M = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad q = \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

It is easy to see that the LCP(M, q) has the solution set $\bar{X} = \{(x_1, x_2) \mid x_1 = 0, x_2 \geq 1\}$. Take the sequence $x^k = (k, 1/k), k = 1, 2, \ldots$, then $Mx^k + q = (1/k - 1, k)$, and

$$\|x^k - \bar{x}(x^k)\|_2 = \|(k, 1/k) - (0, 1)\|_2,$$

$$s(x^k) = \|(-1/k + 1, -k, -k, -1/k, 1 - k + 1)_+\|$$

$$= \|(-1/k+1,\ 2-k)_+\|,$$

$$r(x^k) = \|min\{(k,1/k),\ (1/k-1,\ k)\}\| = \|(1/k-1,\ 1/k)\|.$$

It is easy to see that the distance between x^k and the solution set goes to the infinity, but both $s(x^k)+s^{\frac{1}{2}}(x^k)$ and $r(x^k)+s(x^k)$ remain bounded as k goes to the infinity.

Both residuals $s(x)+s(x)^{\frac{1}{2}}$ and $r(x)+s(x)$ are global error bounds for LCP(M, q) with positive semi-definite M. By Theorem 2.2.1, $r(x)+s(x)$ is a global error bound for R_0-matrices. The following theorem proves that this is also true for $s(x)+s(x)^{\frac{1}{2}}$. Thus this theorem gives another class of matrices for which the error bound of Mangasarian and Shiau ([MaS86]) holds besides the semidefinite case.

Theorem 2.3.3 *Let $M \in I\!R^{n\times n}$ be an R_0-matrix. Then there exists a positive σ such that*

$$\|x-\bar{x}(x)\| \le \sigma(s(x)+s(x)^{\frac{1}{2}}), \tag{2.3.4}$$

where $\bar{x}(x)$ is a closest solution to x under the norm $\|\cdot\|$.

Proof. Assume that the theorem is false. Then for each integer k, there exists an x^k such that (2.3.4) is violated, i.e.

$$\|x^k-\bar{x}(x^k)\| > k(s(x^k)+s(x^k)^{\frac{1}{2}}),$$

where $\bar{x}(x^k)$ is a closest solution of LCP(M, q) to x^k under the norm $\|\cdot\|$.

Case 1. $\|x^k\|$ is unbounded. Without loss of generality, let

$$\lim_{k\to\infty}\frac{x^k}{\|x^k\|} = s.$$

Note that $\|s\| = 1$. By taking a fixed solution $\bar{x}$ as $\bar{x}(x^k)$, it follows that

$$\|x^k - \bar{x}\| \geq \|x^k - \bar{x}(x^k)\| > k(s(x^k) + s(x^k)^{\frac{1}{2}}). \tag{2.3.5}$$

Dividing by $\|x^k\|$ and letting k go to infinity give

$$\begin{aligned} 1 &= \lim_{k\to\infty} \frac{\|x^k - \bar{x}\|}{\|x^k\|} \\ &\geq \lim_{k\to\infty} k(\frac{s(x^k) + s^{\frac{1}{2}}(x^k)}{\|x^k\|}) \end{aligned} \tag{2.3.6}$$

Therefore

$$\lim_{k\to\infty} \frac{s(x^k) + s^{\frac{1}{2}}(x^k)}{\|x^k\|} = 0,$$

otherwise (2.3.6) would be violated. By using the definition of $s(x)$ and taking the limit of the above expression, it follows that

$$\|(-s, -Ms, \lim_{k\to\infty} \frac{x^k(Mx^k + q)}{\|x^k\|})_+\| + (sMs)_+^{\frac{1}{2}} = 0.$$

Therefore $s \geq 0,\ Ms \geq 0,\ sMs = 0$. This contradicts the assumption that M is an R_0-matrix.

Case 2. $\|x^k\|$ is bounded. The left hand side of (2.3.5) is finite when k goes to infinity, thus $s(x^k)$ goes to zero. Without loss of generality, let $\{x^k\}$ converge to x^* and $s(x^*) = 0$. Therefore x^* is a solution. On the other hand, since $r(x)$ is a local error bound for each LCP(M, q) by Lemma 2.2.2, there exist positive K, ϵ and τ such that when $k > K$, $r(x^k) \leq \epsilon$ holds. Therefore for $k > K$

$$\|x^k - \bar{x}(x^k)\| \leq \tau r(x^k). \tag{2.3.7}$$

Let $I = \{x \mid x_i \geq (Mx + q)_i\}$. We now estimate $r(x)$ as follows

$$r(x) = \|x - [x - Mx - q]_+\|_1$$

$$
\begin{aligned}
&= \sum_{i\in I}|(Mx+q)_i| + \sum_{i\notin I}|x_i| \\
&= \sum_{i\in I,(Mx+q)_i\geq 0}|(Mx+q)_i| + \sum_{i\in I,(Mx+q)_i<0}|(Mx+q)_i| \\
&\quad + \sum_{i\notin I,x_i\geq 0}|x_i| + \sum_{i\notin I,x_i<0}|x_i| \\
&\leq \sum_{i\in I,(Mx+q)_i\geq 0}|(Mx+q)_i| + \sum_{i\notin I,x_i\geq 0}|x_i| + \|(-x,\ -Mx-q)_+\|_1 \\
&\leq n^{\frac{1}{2}}[\sum_{i\in I,(Mx+q)_i\geq 0}(Mx+q)_i^2 + \sum_{i\notin I,x_i\geq 0}x_i^2]^{\frac{1}{2}} + \|(-x,\ -Mx-q)_+\|_1 \\
&\leq n^{\frac{1}{2}}[\sum_{i\in I,(Mx+q)_i\geq 0}x_i(Mx+q)_i + \sum_{i\notin I,x_i\geq 0}x_i(Mx+q)_i]^{\frac{1}{2}} \\
&\quad +\|(-x,\ -Mx-q)_+\|_1 \\
&\leq n^{\frac{1}{2}}[\sum_{x_i(Mx+q)_i\geq 0}^{n}x_i(Mx+q)_i]^{\frac{1}{2}} + \|(-x,\ -Mx-q)_+\|_1 \\
&= n^{\frac{1}{2}}[\sum_{x_i(Mx+q)_i\geq 0}x_i(Mx+q)_i + \sum_{x_i(Mx+q)_i<0}x_i(Mx+q)_i \\
&\quad - \sum_{x_i(Mx+q)_i<0}x_i(Mx+q)_i]^{\frac{1}{2}} + \|(-x,\ -Mx-q)_+\|_1 \\
&= n^{\frac{1}{2}}[x(Mx+q) - \sum_{x_i(Mx+q)_i<0}x_i(Mx+q)_i]^{\frac{1}{2}} + \|(-x,\ -Mx-q)_+\|_1 \\
&\leq n^{\frac{1}{2}}[(x(Mx+q))_+ - \sum_{x_i(Mx+q)_i<0}x_i(Mx+q)_i]^{\frac{1}{2}} + \|(-x,\ -Mx-q)_+\|_1 \\
&\leq n^{\frac{1}{2}}[(x(Mx+q))_+^{\frac{1}{2}} + (\sum_{x_i(Mx+q)_i<0}|x_i(Mx+q)_i|)^{\frac{1}{2}}] \qquad (2.3.8) \\
&\quad +\|(-x,\ -Mx-q)_+\|_1 \\
&\leq n^{\frac{1}{2}}[(s(x)+s(x)^{\frac{1}{2}}) + (\sum_{x_i(Mx+q)_i<0}|x_i(Mx+q)_i|)^{\frac{1}{2}}]. \qquad (2.3.9)
\end{aligned}
$$

By combining (2.3.5), (2.3.7) and (2.3.8), it follows that

$$
\begin{aligned}
&k(s(x^k)+s(x^k)^{\frac{1}{2}}) < \|x^k - \bar{x}(x^k)\| \leq \tau r(x^k) \\
&\leq \ \tau n^{\frac{1}{2}}\left\{(s(x^k)+s(x^k)^{\frac{1}{2}}) + (\sum_{x_i^k(Mx^k+q)_i<0}|x_i^k(Mx^k+q)_i|)^{\frac{1}{2}}\right\}. \qquad (2.3.10)
\end{aligned}
$$

But

$$(k - \tau n^{\frac{1}{2}})\|(-x^k, -Mx^k - q)_+\|_1^{\frac{1}{2}} \le (k - \tau n^{\frac{1}{2}})(s(x^k) + s(x^k)^{\frac{1}{2}}) \qquad (2.3.11)$$

Combining (2.3.10) and (2.3.11) gives

$$(k - \tau n^{\frac{1}{2}})\|(-x^k, -Mx^k - q)_+\|_1^{\frac{1}{2}} < \tau n^{\frac{1}{2}}(\sum_{x_i^k(Mx^k+q)_i<0} |x_i^k(Mx^k + q)_i|)^{\frac{1}{2}}.$$

Since x^k and $Mx^k + q$ are bounded, i.e. $|x_i^k| \le N$ and $|(Mx^k + q)_i| \le N$, $i = 1, \ldots, n$ for some fixed $N > 0$, take $(k - \tau n^{\frac{1}{2}}) \ge \tau(nN)^{\frac{1}{2}}$ and the above inequality fails to hold. Therefore, we get the contradiction. **Q.E.D.**

2.4 Comparisons Among Different Error Residuals

In this section, we establish relationships among different error residuals. We note that we have not been able to provide a global error bound for an indefinite matrix M which is not of the type $(1 + \|x\|)e(x)$ where $e(x) = 0$ on the solution set. We exhibit all known error bounds in a useful table format given in Table 1.1. First we define some other error residuals.

Definition 2.4.1 *Let M be in $\mathbb{R}^{n\times n}$. Define the following error residuals for LCP(M, q)*

$$t(x) := \|(-Mx - q, \ -x)_+, \sum_{i=1}^n (x_i(Mx + q)_i)_+\|, \qquad (2.4.1)$$

$$v(x) := \|(-Mx - q, \ -x)_+, \ \sum_{i=1}^n |x_i(Mx + q)_i|\|$$

It is obvious that $s(x) \le t(x) \le v(x)$. In addition we find a relationship between $r(x)$ and $t(x)$ as can be seen from the following proposition.

Proposition 2.4.1 *Let $\|\cdot\|$ denote the 1-norm in the definition of $r(x)$ and $t(x)$. Then for any $x \in I\!R^n$,*

$$r(x) \leq \|(-x,\ -Mx-q)_+\|_1 + [n\sum_{i=1}^{n}(x_i(Mx+q)_i)_+]^{1/2}.$$

Consequently,

$$r(x) \leq n^{1/2}(t(x) + t(x)^{1/2}) \tag{2.4.2}$$

Proof. Let $I = \{i \mid x_i \geq (Mx+q)_i\}$, then

$$\begin{aligned}
r(x) &= \|x - [x - Mx - q]_+\|_1 \\
&= \sum_{i\in I}|(Mx+q)_i| + \sum_{i\notin I}|x_i| \\
&= \sum_{i\in I,(Mx+q)_i\geq 0}|(Mx+q)_i| + \sum_{i\in I,(Mx+q)_i<0}|(Mx+q)_i| \\
&\quad + \sum_{i\notin I,x_i\geq 0}|x_i| + \sum_{i\notin I,x_i<0}|x_i| \\
&\leq \sum_{i\in I,(Mx+q)_i\geq 0}|(Mx+q)_i| + \sum_{i\notin I,x_i\geq 0}|x_i| + \|(-x,\ -Mx-q)_+\|_1 \\
&\leq n^{\frac{1}{2}}[\sum_{i\in I,(Mx+q)_i\geq 0}(Mx+q)_i^2 + \sum_{i\notin I,x_i\geq 0}x_i^2]^{\frac{1}{2}} + \|(-x,\ -Mx-q)_+\|_1 \\
&\leq n^{\frac{1}{2}}[\sum_{i\in I,(Mx+q)_i\geq 0}(x_i(Mx+q)_i)_+ + \sum_{i\notin I,x_i\geq 0}(x_i(Mx+q)_i)_+]^{\frac{1}{2}} \\
&\quad + \|(-x,\ -Mx-q)_+\|_1 \\
&\leq n^{\frac{1}{2}}[\sum_{i=1}^{n}(x_i(Mx+q)_i)_+]^{\frac{1}{2}} + \|(-x,\ -Mx-q)_+\|_1.
\end{aligned}$$

Q.E.D.

Remark. The converse of Proposition 2.4.1 is not true, i.e. the residual $r(x)$ cannot bound the residual $t(x) + t(x)^{\frac{1}{2}}$ as the following example shows. In fact, the following example shows that $r(x)$ cannot bound even the smaller residual $s(x) + s(x)^{\frac{1}{2}}$. Let

$$M = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad q = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

It is easy to see that $\bar{x} = (0,\ 0)$ is the unique solution for the LCP. Now let the sequence $x^k = (-\frac{1}{k},\ 0),\ k = 1, 2, \ldots$. Then $Mx^k + q = x^k$ and under the 1-norm

$$r(x^k) = \|x^k - (x^k - Mx^k - q)_+\|_1 = \frac{1}{k},$$

$$s(x^k) = \|(-Mx^k - q,\ -x^k,\ x^k(Mx^k + q))_+\|_1 = 2\frac{1}{k} + \frac{1}{k^2}.$$

It follows that

$$\begin{aligned} &\lim_{k\to\infty} \frac{s(x^k) + s(x^k)^{\frac{1}{2}}}{r(x^k)} \\ &= \lim_{k\to\infty} 2 + \frac{1}{k} + (2k+1)^{\frac{1}{2}} \\ &= +\infty \end{aligned}$$

Remark. Although from Proposition 2.4.1, $r(x)$ can be bounded by $t(x) + t(x)^{\frac{1}{2}}$, it can not be bounded by the smaller residual $s(x) + s(x)^{\frac{1}{2}}$ as the following example implies. Let

$$M = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad q = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The solution set $\bar{X}$ of LCP(M, q) is $\{x \mid x_1 = x_3 = 0, x_2 \geq 0 \text{ or } x_2 = x_3 = 0, x_1 \geq 0\}$. Take $x^k = (-k^{-4}, k^2, k^{-1})$, then $Mx^k + q = (k^2, 0, k^{-1})$ and

$$r(x^k) = \|min\{x^k, Mx^k + q\}\| = \|(-k^{-4}, 0, k^{-1})\|$$

but

$$s(x^k) = \|(-Mx^k - q, -x^k, x^k(Mx^k + q)\| = k^{-4}$$

Therefore there does not exist a constant τ such that

$$r(x^k) \leq \tau(s(x^k) + s(x^k)^{\frac{1}{2}}).$$

This example even gives a further implication regarding $s(x^k)+s(x^k)^{\frac{1}{2}}$, that is, it is not a local error bound for LCP(M, q). This is so because

$$\|x^k-\bar{x}(x^k)\|_2=\|(-k^{-4},0,k^{-1}\|_2$$

which goes to zero slower than $s(x^k)+s(x^k)^{\frac{1}{2}}$. In addition, since $r(x^k)$ goes to zero, we know that even locally $s(x^k)+s(x^k)^{\frac{1}{2}}$ cannot bound $r(x^k)$.

Remark. There is another interesting comparison between $r(x)$ and $s(x)+s(x)^{\frac{1}{2}}$, that is, if $r(x)$ goes to zero, then the distance between x and the solution set of an LCP(M, q) goes to zero. This is because $r(x)$ is a local error bound. However, as the following example indicates, although $s(x)+s(x)^{\frac{1}{2}}$ goes to zero, the distance goes to infinity. Let

$$M=\begin{bmatrix}0 & -1\\ 0 & 1\end{bmatrix},\qquad q=\begin{bmatrix}1\\ 1\end{bmatrix}.$$

The LCP(M, q) has unique solution $\bar{x}=(0,0)$. Take $x^k=(k(1+\frac{1}{k})(2+\frac{1}{k}),1+\frac{1}{k})$, then $Mx^k+q=(-1-\frac{1}{k}+1,1+\frac{1}{k}+1)=(-\frac{1}{k},2+\frac{1}{k})$,

$$\begin{aligned}
\|x^k-\bar{x}\| &\to \infty \qquad (k\to\infty).\\
r(x^k) &= \|\min\{x^k,Mx^k+q\}\|\\
&= \|(-\frac{1}{k},1+\frac{1}{k})\|\\
&\to 1 \qquad (k\to\infty).\\
s(x^k) &= \|(-Mx^k-q,-x^k,x^k(Mx^k+q))_+\|\\
&= \|(\frac{1}{k},-2-\frac{1}{k},-k(1+\frac{1}{k})(2+\frac{1}{k}),-1-\frac{1}{k},\\
&\quad -(1+\frac{1}{k})(2+\frac{1}{k})+(1+\frac{1}{2})(2+\frac{1}{k}))_+\|\\
&= \frac{1}{k}\\
&\to 0 \qquad (k\to\infty).
\end{aligned}$$

Now we consider another residual defined as $p(x) := \|(-x,\ -Mx-q)_+\| + [n\sum_{i=1}^n (x_i(Mx+q)_i)_+]^{\frac{1}{2}}$. Under the 1-norm it follows from the Lemma 2.4.1 that

$$r(x) \le p(x) \le n^{\frac{1}{2}}(t(x) + t(x)^{\frac{1}{2}}). \qquad (2.4.3)$$

Hence $p(x)$ is a residual for any LCP(M, q) and it is a local error bound since it bounds $r(x)$. However it does not provide a global error bound for an LCP (2.1.1) for a positive semi-definite M as can be seen from the following example. Let

$$M = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \qquad q = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

M is positive semi-definite since $M + M^T = 0$ and the LCP has the unique solution $\bar{x} = (0, 0)$. Let $x^k = (0,\ k)$. Under 1-norm,

$$\begin{aligned} p(x^k) &= \|(-x^k,\ -Mx^k - q)_+\| + [n\sum_{i=1}^n (x_i^k(Mx^k+q)_i)_+]^{\frac{1}{2}} \\ &= (n(0,\ k)(k+1,\ 1))^{\frac{1}{2}} \\ &= \sqrt{nk}. \end{aligned}$$

It follows that

$$\begin{aligned} &\lim_{k\to\infty} \frac{\|x^k - \bar{x}\|}{p(x^k)} \\ &= \lim_{k\to\infty} \frac{k}{\sqrt{nk}} \\ &= +\infty \end{aligned}$$

We also point out that we cannot reverse the inequality sign in (2.4.3) even if we are permitted to multiply $r(x)$ and $p(x)$ by constants.

So far, we have studied several error residuals such as $p(x), r(x), s(x)$ and $t(x)$. We have explored a number of relationships among them and by combining them, we have obtained new error bounds such as $r(x) + s(x)$ and $t(x) + t(x)^{\frac{1}{2}}$. Now we further ask: What are necessary conditions for some residual to be an error bound? In other words: How big should a residual be to become an error bound? The following theorem is simple, but essential for our answer.

Theorem 2.4.1 *Let $M \in \mathbb{R}^{n\times n}$ and $c = 2 + \|M\|$, then*

$$r(x) \leq c\|x - \bar{x}(x)\|, \tag{2.4.4}$$

where $\bar{x}(x)$ is a closest solution to x under norm $\|\cdot\|$.

Proof. By the definition of $r(x)$, it follows that

$$\begin{aligned} r(x) &= \|x - (x - Mx - q)_+\| \\ &= \|x - (x - Mx - q)_+ - \bar{x}(x) + (\bar{x}(x) - M\bar{x}(x) - q)_+\| \\ &\leq \|x - \bar{x}(x)\| + \|(x - Mx - q)_+ - (\bar{x}(x) - M\bar{x}(x) - q)_+\| \\ &\leq \|x - \bar{x}(x)\| + \|(x - Mx - q) - (\bar{x}(x) - M\bar{x}(x) - q)\| \\ &\leq \|x - \bar{x}(x)\| + \|x - \bar{x}(x)\| + \|Mx - M\bar{x}(x)\| \\ &\leq c\|x - \bar{x}(x)\| \end{aligned}$$

Q.E.D.

Theorem 2.4.1 implies that the order of the distance from any point x to the solution set $\bar{X}$ of any LCP(M, q) is at least as big as $r(x)$. Therefore, in order to be an error bound, a residual must bound $r(x)$. In addition, since $r(x)$ is a local bound by Lemma 2.2.2, locally $r(x)$ is equivalent to the distance to the solution set. So, this distance is precisely characterized by $r(x)$ at least locally. This is a very useful result because $r(x)$, which is a computable quantity, can be used as an alterative for the non-computable distance to the solution set. For other types of residuals we discussed, Theorem 2.4.1 fails to hold unfortunately.

We point out that all these error residuals we have studied are not good enough to globally bound the distance from any point x to the solution set $\bar{X}$ of any LCP(M, q). Consider the following example. Let

$$M = \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}, \qquad q = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

It is easy to see that the LCP(M, q) has the unique solution $\bar{x} = (0,0)$. Take the sequence $x^k = (k, 1-1/k), k = 1,2,...,$ then $Mx^k + q = (1/k, 2-1/k)$ and

$$\begin{aligned} & \lim_{k\to\infty} \frac{\|x-\bar{x}\|}{v(x)+v(x)^{1/2}} \\ = \; & \lim_{k\to\infty} \frac{\|(k, 1-\frac{1}{k})\|}{\|(0,0,1+(1-\frac{1}{k})(2-\frac{1}{k}))\| + \|(0,0,1+(1-\frac{1}{k})(2-\frac{1}{k}))\|^{1/2}} \\ = \; & \infty. \end{aligned}$$

This example shows that Theorem 2.3.1 is false for an indefinite matrix M. It is not clear to what extent we can extend Theorem 2.3.1. For example, we are not able to show that it holds even for an indefinite symmetric matrix M, nor can we give a counter-example.

Although all these error residuals fail to provide a global error for an arbitrary LCP, there is a way to globalize a local error bound as can be seen in the following theorem.

Theorem 2.4.2 *Let $M \in \mathbb{R}^{n\times n}$ and $l(x)$ be any local error bound. Then, there exists a positive τ such that*

$$\|x - \bar{x}(x)\| \le \tau(1 + \|x\|)l(x), \tag{2.4.5}$$

where $\bar{x}(x)$ is a closest solution from x to the solution set of LCP(M, q)

Proof. Assume that the theorem is false. then for each k, there exists an x^k such that (2.4.5) is violated, i.e.

$$\|x^k - \bar{x}(x^k)\| \geq k(1 + \|x^k\|)l(x^k),$$

where $\bar{x}(x^k)$ is a closest solution from x^k to the solution set of LCP(M, q). Then for a fixed solution $\bar{x}$, we have

$$\|x^k - \bar{x}\| \geq \|x^k - \bar{x}(x^k)\| \geq k(1 + \|x^k\|)l(x^k). \tag{2.4.6}$$

Since $l(x)$ is a local error bound by Lemma 2.2.2, there exist $K > 0$ and $\epsilon > 0$ such that $l(x^k) > \epsilon$, for $k > K$ (See proof of Theorem 2.2.1). Hence the right hand side above goes to infinity as k goes to infinity and so does the left hand side since it is bigger. Therefore, $\|x^k\|$ goes to infinity. Without loss of generality, let

$$\lim_{k\to\infty} \frac{x^k}{\|x^k\|} = s$$

and $s \neq 0$ since $\|s\| = 1$. Divide both sides of (2.4.6) by $\|x^k\|$ and let k goes to infinity, then it follows that

$$\begin{aligned} 1 &= \lim_{k\to\infty} \frac{\|x^k - \bar{x}\|}{\|x^k\|} \\ &\geq \lim_{k\to\infty} k\frac{(1 + \|x^k\|)l(x^k)}{\|x^k\|} \\ &= \lim_{k\to\infty} kl(x^k) \\ &= \lim_{k\to\infty} k\epsilon. \end{aligned}$$

We get contradiction since the left hand side is finite. Therefore, Theorem 2.4.2 holds. **Q.E.D.**

Remark. Note that in (2.4.5) if x is far away from the origin, then the error bound value increases, which could have nothing to do with the actual distance between x and the solution set of an LCP. This is a major drawback of this type of global error bounds. There are similar cases in [Ma91].

Now, we summarize our error bound relationships in Table1.1 where the following definition have been used:

$$\begin{aligned} r(x) &:= \|x - (x - Mx - q)_+\| \\ s(x) &:= \|(-Mx - q, -x, x(Mx + q))_+\| \\ t(x) &:= \|(-Mx - q, -x)_+, \sum_{i=1}^{n}(x_i(Mx + q)_i)_+\| \\ \text{psd} &:= \text{positive semi-definite } M \\ \text{pd} &:= \text{positive definite } M \\ R_0 &:= \text{M such that zero is only solution for LCP(M, 0).} \end{aligned}$$

We also note that there exist positive τ_1, τ_2 and c such that:

$$\begin{aligned} r(x) &\leq c\|x - \bar{x}(x)\| \\ s(x) + s(x)^{\frac{1}{2}} &\leq \tau_1(t(x) + t(x)^{\frac{1}{2}}) \\ r(x) + s(x)^{\frac{1}{2}} &\leq \tau_2(t(x) + t(x)^{\frac{1}{2}}) \\ t(x) &\leq v(x). \end{aligned}$$

Table 1.1. Validity of Various Residuals as Local and Global Error Bounds for LCP

$M \in \mathbb{R}^{n\times n}$	$r(x)$	$s(x) + s(x)^{\frac{1}{2}}$	$r(x) + s(x)$	$t(x) + t(x)^{\frac{1}{2}}$
arbitrary	local	not local	local	local
psd	local	global	global	global
R_0	global	global	global	global
pd	global	global	global	global

It is also interesting to note that the residual $r(x)+s(x)$, which can be thought of as the average of $r(x)$ and $s(x)$, covers most cases without recourse to the irrational square root residual. This in a certain sense can be thought of as the best residual.

Chapter 3

Error Bounds for the Strongly Monotone Nonlinear Complementarity Problem (NCP)

3.1 Introduction

In this chapter, we search for new error bounds for nonlinear complementarity problems. In general it is very difficult to establish even a local error bound for nonlinear complementarity problems. In fact all error residuals used for the linear complementarity problem in Chapter 2 fail to hold even as local error bounds for the nonlinear complementarity problem. The only case where some of these residuals work is when the nonlinear complementarity problems is strongly monotone. We establish this result in the next section.

Consider the nonlinear complementarity problem (NCP for short) of finding an x in $\mathbb{R}^n$ such that

$$F(x) \geq 0, \quad x \geq 0, \quad xF(x) = 0, \tag{3.1.1}$$

where $F(x)$ is a function from $\mathbb{R}^n$ to $\mathbb{R}^n$. The LCP (2.1.1) obtains when $F(x) = Mx + q$. We now define residuals for the NCP that correspond to ones defined for the LCP as follows:

$$\begin{aligned} r(x) &:= \|x - (x - F(x))_+\|, \\ s(x) &:= \|(-x, \ -F(x), \ xF(x))_+\|, \\ t(x) &:= \|(-x, \ -F(x), \ \sum_{i=1}^{n}(x_iF_i(x))_+\|, \\ v(x) &:= \|(-x, \ -F(x), \ \sum_{i=1}^{n}|x_iF_i(x)|\| \end{aligned} \tag{3.1.2}$$

Note that $\bar{x}$ is a solution of (3.1.1) if and only if any of the residuals above is equal to zero. An NCP is strongly monotone if there exists a constant $c > 0$ such that for any $x, \ y \in \mathbb{R}^n$

$$(F(x) - F(y))(x - y) \geq c\|x - y\|^2 \tag{3.1.3}$$

3.2 A New Error Bound for Strongly Monotone NCP

In this section, we establish that the residual $r(x)$ is a global error bound for the strongly monotone NCP. Here is the main theorem.

Theorem 3.2.1 *Let the NCP (3.1.1) be strongly monotone with constant $c > 0$. Assume that $F(x)$ is Lipschitz continuous with constant $L > 0$, that is*

$$\|F(x) - F(y)\| \leq L\|x - y\|, \quad \forall x, \ y \in \mathbb{R}^n.$$

Then $\forall x \in I\!R^n$

$$\|x - \bar{x}\| \leq \frac{1+2L}{c}\|r(x)\|, \tag{3.2.1}$$

where $\bar{x}$ *is the unique solution of the NCP.*

Proof. For a **fixed** x, the point $p(x) := (x - F(x))_+$ is a solution of the related LCP

$$\bar{F}(p) := p - x + F(x) \geq 0, \quad p \geq 0, \quad p\bar{F}(p) = 0.$$

It follows that

$$\begin{aligned}
&(p(x) - \bar{x})(\bar{F}(p(x)) - F(\bar{x})) \\
&= p(x)\bar{F}(p(x)) - p(x)F(\bar{x}) - \bar{x}\bar{F}(p(x)) + \bar{x}F(\bar{x}) \\
&= -p(x)F(\bar{x}) - \bar{x}\bar{F}(p(x)) \\
&\leq 0.
\end{aligned}$$

By the definition $\bar{F}(p(x)) = p(x) - x + F(x)$, it follows from the inequality above that

$$\begin{aligned}
0 &\geq (p(x) - \bar{x})(p(x) - x) + (p(x) - \bar{x})(F(x) - F(\bar{x})) \\
&= (p(x) - \bar{x})(p(x) - x) + (p(x) - x + x - \bar{x})(F(x) - F(\bar{x})) \\
&= (p(x) - \bar{x})(p(x) - x) + (p(x) - x)(F(x) - F(\bar{x})) \\
&\quad +(x - \bar{x})(F(x) - F(\bar{x})) \\
&= (p(x) - x)(p(x) - \bar{x} + F(x) - F(\bar{x})) + (x - \bar{x})(F(x) - F(\bar{x})) \\
&\geq (p(x) - x)(p(x) - \bar{x} + F(x) - F(\bar{x})) + c\|x - \bar{x}\|^2. \qquad (3.2.2)
\end{aligned}$$

Since $\|F(x) - F(\bar{x})\| \leq L\|x - \bar{x}\|$, it follows that

$$\|p(x) - \bar{x}\| = \|p(x) - p(\bar{x})\|$$

$$\begin{aligned} &= \|(x - F(x))_+ - (\bar{x} - F(\bar{x}))_+\| \\ &\leq \|x - F(x) - (\bar{x} - F(\bar{x}))\| \\ &\leq \|x - \bar{x}\| + \|F(x) - F(\bar{x}))\| \\ &\leq \|x - \bar{x}\| + L\|x - \bar{x}\| \\ &\leq (1 + L)\|x - \bar{x}\| \end{aligned} \tag{3.2.3}$$

By applying Cauchy-Schwartz to (3.2.2), we have

$$\begin{aligned} c\|x - \bar{x}\|^2 &\leq \|x - p(x)\|(\|p(x) - \bar{x}\| + \|F(x) - F(\bar{x})\|) \\ &\leq r(x)((1 + L)\|x - \bar{x}\| + L\|x - \bar{x}\|) \quad \text{By (3.2.3)} \\ &\leq (1 + 2L)r(x)\|x - \bar{x}\| \end{aligned}$$

Therefore,

$$\|x - \bar{x}\| \leq \frac{1 + 2L}{c} r(x).$$

Q.E.D.

We have a similar relationship between $r(x)$ and $t(x)$ to that for the the corresponding residuals for the LCP, Proposition 2.4.1.

Proposition 3.2.1 *Let $\|\cdot\|$ denote the 1-norm in the definition of $r(x)$ and $t(x)$. Then for any $x \in \mathbb{R}^n$,*

$$r(x) \leq \|(-x, -F(x))_+\|_1 + [n \sum_{i=1}^{n} (x_i F_i(x))_+]^{\frac{1}{2}}.$$

Consequently,

$$r(x) \leq n^{1/2}(t(x) + t(x)^{1/2}) \tag{3.2.4}$$

Proof. The proof is similar to that of Lemma 2.4.1.

Theorem 3.2.2 *Under the assumptions of Theorem 3.2.1, there exists some constant $\sigma > 0$ such that*

$$\|x - \bar{x}\| \leq \sigma(t(x) + t(x)^{\frac{1}{2}}).$$

where $\bar{x}$ is the unique solution of the NCP.

Proof. The proof follows directly from Theorem 3.2.1 and Lemma 3.2.1.

An important feature of the residual $r(x)$ is that it is also a lower bound to the distance between any point x and the solution set of the NCP (3.1.1), provided that $F(x)$ is Lipschitz continuous. This is characterized in the following theorem.

Theorem 3.2.3 *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be Lipschitz continuous with constant L. Assume that the NCP (3.1.1) has a solution. Then for any solution $\bar{x}$ of the NCP (3.1.1),*

$$r(x) \leq (2 + L)\|x - \bar{x}\|. \tag{3.2.5}$$

where $\|\cdot\|$ is a monotonic norm on $\mathbb{R}^n$.

Proof. By the definition of $r(x)$, it follows

$$\begin{aligned} r(x) &\leq \|x - (x - F(x))_+\| \\ &= \|x - (x - F(x))_+ - \bar{x} + (\bar{x} - F(\bar{x}))_+\| \\ &\leq \|x - \bar{x}\| + \|(x - F(x))_+ - (\bar{x} - F(\bar{x})_+\| \\ &\leq \|x - \bar{x}\| + \|x - \bar{x}\| + \|F(x) - F(\bar{x})\| \\ &\leq (2 + L)\|x - \bar{x}\|. \end{aligned}$$

Q.E.D.

Remark. By Theorems 3.2.1 and 3.2.3, we have that the distance $\|x - \bar{x}\|$ is equivalent to $r(x)$ in the sense that $r(x)$ can bound $\|x - \bar{x}\|$ from both above and below. In other words, $r(x)$ is a very good characterization of the distance $\|x - \bar{x}\|$. More importantly, it is easy to compute for any x. So, it can be used to measure whether an approximate solution is a good solution and thus $r(x)$ can serve as a termination criteria for an algorithm.

It is interesting that without the Lipschitz continuity of $F(x)$, the residual $s(x) + s(x)^{\frac{1}{2}}$ provides a global error bound for strongly monotone nonlinear complementarity problems. But this is not the case for $r(x)$ in Theorem 3.2.1. The following theorem gives this result precisely which extends the result of [MaS86] to the NCP for the strongly monotone case.

Theorem 3.2.4 *Let $F : I\!R^n \to I\!R^n$ be strongly monotone. Assume that the NCP (3.1.1) has a solution $\bar{x}$ which is unique. Then there exists a positive λ such that*

$$\|x - \bar{x}\| \le \lambda(s(x) + s(x)^{\frac{1}{2}}). \tag{3.2.6}$$

where the norm is a monotonic norm.

Proof. Assume that the theorem is false. Then for each integer k, there exists an x^k such that

$$\|x^k - \bar{x}\| > k(s(x^k) + s(x^k)^{\frac{1}{2}}). \tag{3.2.7}$$

Case 1. Let $\{x^k\}$ be unbounded. Then dividing the above inequality by $\|x^k\|$ and letting k go to infinity gives

$$\begin{aligned} 1 &= \lim_{k\to\infty} \frac{\|x^k - \bar{x}\|}{\|x^k\|} \\ &\ge \lim_{k\to\infty} \frac{k(s(x^k) + s(x^k)^{\frac{1}{2}})}{\|x^k\|} \\ &\ge \lim_{k\to\infty} k\frac{s(x^k)}{\|x^k\|} \\ &= \lim_{k\to\infty} k\|(-\frac{x^k}{\|x^k\|}, -\frac{F(x^k)}{\|x^k\|}, \frac{x^k F(x^k)}{\|x^k\|})_+\|. \end{aligned} \tag{3.2.8}$$

Hence

$$1 \geq \lim_{k\to\infty} k\|(-\frac{x^k}{\|x^k\|}, -\frac{F(x^k)}{\|x^k\|}, \frac{x^kF(x^k)}{\|x^k\|})_+\|. \tag{3.2.9}$$

Therefore, $\lim_{k\to\infty} -\frac{x^k}{\|x^k\|} \leq 0, \lim_{k\to\infty} -\frac{F(x^k)}{\|x^k\|} \leq 0$. But since $\bar{x} \geq 0$ and $F(\bar{x}) \geq 0$, we have

$$\begin{aligned}
\lim_{k\to\infty} \frac{x^kF(x^k)}{\|x^k\|} &\geq \lim_{k\to\infty} \frac{x^kF(x^k) + \bar{x}F(\bar{x}) - x^kF(\bar{x}) - \bar{x}F(x^k)}{\|x^k\|} \\
&\geq \frac{(F(x^k) - F(\bar{x}))(x^k - \bar{x})}{\|x^k\|} \\
&\geq \frac{c\|x^k - \bar{x}\|^2}{\|x^k\|} \\
&= \infty.
\end{aligned}$$

This contradicts (3.2.9)

Case 2. Let $\|x^k\|$ be bounded. Without loss of generality, let $\{x^k\}$ converge to $\tilde{x}$. Since the left hand side of (3.2.7) is bounded, so $s(x^k)$ goes to zero. Therefore $s(\tilde{x}) = 0$, and hence $\tilde{x}$ is a solution of the NCP. Since the NCP has a unique solution, $\bar{x} = \tilde{x}$. Again from (3.2.7),

$$\begin{aligned}
\|x^k - \bar{x}\| &> ks(x^k)^{\frac{1}{2}} \\
&= k\|(-x^k, -F(x^k), x^kF(x^k))_+\|^{\frac{1}{2}} \\
&= k\|(-x^k, -F(x^k), x^kF(x^k) + \bar{x}F(\bar{x}) - \bar{x}F(x^k) - x^kF(\bar{x}) \\
&\quad +\bar{x}F(x^k) + x^kF(\bar{x}))_+\|^{\frac{1}{2}} \\
&= k\|(-x^k, -F(x^k), (F(x^k) - F(\bar{x}))(x^k - \bar{x}) + \bar{x}F(x^k) + x^kF(\bar{x}))_+\|^{\frac{1}{2}} \\
&\geq k\|(-x^k, -F(x^k), c\|x^k - \bar{x}\|^2 + \bar{x}F(x^k) + x^kF(\bar{x}))_+\|^{\frac{1}{2}}
\end{aligned}$$

Hence

$$\|x^k - \bar{x}\| > k\|(-x^k, -F(x^k), c\|x^k - \bar{x}\|^2 + \bar{x}F(x^k) + x^kF(\bar{x}))_+\|^{\frac{1}{2}}. \tag{3.2.10}$$

We prove that this inequality cannot hold when k is sufficiently large, and hence we have a contradiction. In fact, if $\bar{x}F(x^k) + x^kF(\bar{x}))_+ < -\frac{c}{2}||x^k - \bar{x}||^2$, then

$$\begin{aligned}
\frac{c}{2}||x^k - \bar{x}||^2 &< -\bar{x}F(x^k) - x^kF(\bar{x}))_+ \\
&\leq (-x^k)_+F(\bar{x}) + (-F(x^k))_+\bar{x} \\
&\leq ||(-x^k, -F(x^k))_+||||\bar{x}, F(\bar{x})||\prime.
\end{aligned}$$

where $||\cdot||\prime$ is the dual norm to $||\cdot||$. It is easy to see that $||\bar{x}, F(\bar{x})||\prime > 0$. Hence from (3.2.10)

$$\begin{aligned}
||x^k - \bar{x}||^2 &> k^2||(-x^k, -F(x^k))_+|| \\
&\geq \frac{k^2c||x^k - \bar{x}||^2}{2||\bar{x}, F(\bar{x})||}.
\end{aligned}$$

Hence

$$1 > \frac{k^2c}{2||\bar{x}, F(\bar{x})||}.$$

This inequality fails to hold for k sufficiently large.

Suppose now that $\bar{x}F(x^k) + x^kF(\bar{x}))_+ \geq -\frac{c}{2}||x^k - \bar{x}||^2$, then by (3.2.10)

$$\begin{aligned}
||x^k - \bar{x}|| &> k||(-x^k, -F(x^k), c||x^k - \bar{x}||^2 + \bar{x}F(x^k) + x^kF(\bar{x}))_+||^{\frac{1}{2}} \\
&\geq k||(c||x^k - \bar{x}||^2 + \bar{x}F(x^k) + x^kF(\bar{x}))_+||^{\frac{1}{2}} \\
&\geq k||(c||x^k - \bar{x}||^2 - \frac{c}{2}||x^k - \bar{x}||^2)_+||^{\frac{1}{2}} \\
&= k\sqrt{\frac{c}{2}}||x^k - \bar{x}||.
\end{aligned}$$

Hence when $k \geq \sqrt{\frac{2}{c}}$, this inequality fails to hold. Therefore we have a contradiction. **Q.E.D.**

Chapter 4

The Implicit Lagrangians as a New Error Bound for the Linear and Nonlinear Complementarity Problem

4.1 The Implicit Lagrangian $M(x, \alpha)$

Recently Mangasarian and Solodov [MaS92] established the following interesting relation between each NCP (3.1.1) and the following implicit Lagrangian function for (3.1.1)

$$\begin{aligned} M(x,\alpha) \;:=\; & 2\alpha x F(x) + \|(-\alpha F(x) + x)_+\|^2 - \|x\|^2 \\ & + \|(-\alpha x + F(x))_+\|^2 - \|F(x)\|^2, \end{aligned} \tag{4.1.1}$$

where $\|\cdot\|$ denotes the 2-norm in this chapter and α is some fixed positive real number. In particular, for each $\alpha > 1$, they proved that the implicit Lagrangian function $M(x, \alpha) \geq 0$, $\forall x \in R^n$ and any point x solves NCP (3.1.1) if and only if $M(x, \alpha) = 0$. Therefore we can think of $M(x, \alpha)$ as a residual. More recently, Luo [LMRS92] proved that if $F(x)$ is linear, that is $F(x) = Mx + q$, then $(M(x, \alpha))^{\frac{1}{2}}$ does indeed provide a local error bound for the distance between any point x and the solution set of the LCP (2.1.1), that is, there exist some constants $\kappa > 0$ and $\delta > 0$ such that

$$\|x - \bar{x}(x)\| \leq \kappa (M(x, \alpha))^{\frac{1}{2}}, \quad \forall x \text{ whenever } M(x, \alpha) \leq \delta, \tag{4.1.2}$$

where $\bar{x}(x)$ is an orthogonal projection of x on the solution set of the LCP (2.1.1).

In this chapter we consider all residuals defined in (3.1.2) including $M(x, \alpha)$. We want to investigate the relationships between the residual $M(x, \alpha)$ and those known residuals such as $r(x)$, $s(x)$, $t(x)$ and $v(x)$. As a result, we can establish error bound results for the residual $(M(x, \alpha))^{\frac{1}{2}}$. In the next section we prove an equivalence relationship between the residuals $(M(x, \alpha))^{\frac{1}{2}}$ and $r(x)$. In the last section of this chapter, the restricted implicit Lagrangian $N(x, \alpha)$ ([Fuk92], [MaS92]) is compared with the known error residuals. Consequently, error bound results for $N(x, \alpha)$ are obtained.

4.2 An Equivalence Relation between $r(x)$ and $(M(x, \alpha))^{\frac{1}{2}}$

We begin with a theorem that establishes a precise inequality relation between $r(x)$ and $((M(x, \alpha))^{\frac{1}{2}}$.

Theorem 4.2.1 *Let $F : I\!R^n \to I\!R^n$ be any function. Then for each $\alpha > 1$, the following relation holds.*

$$2(\alpha - 1) r(x)^2 \leq M(x, \alpha) \leq 2\alpha(\alpha - 1) r(x)^2, \tag{4.2.1}$$

Proof. Let

$$M_i(x,\alpha) := 2\alpha x_i F_i(x) + (-\alpha F_i(x) + x_i)_+^2 - x_i^2 + (-\alpha x_i + F_i(x))_+^2 - F_i(x)^2$$

and

$$r_i(x) := (x - (x - F(x))_+)_i = min\{x_i,\ F_i(x)\}.$$

Then

$$M(x,\ \alpha) = \sum_{i=1}^{n} M_i(x,\alpha)$$

and

$$r(x)^2 = \sum_{i=1}^{n} r_i(x)^2.$$

So, in order to prove (4.2.1), it is sufficient to show that

$$2(\alpha-1)r_i(x)^2 \le M_i(x,\alpha) \le 2\alpha(\alpha-1)r_i(x)^2.$$

For convenience, let $I = \{i \mid x_i \ge \alpha F_i(x)\}$ and $J = \{j \mid F_j(x) \ge \alpha x_j\}$ while $\bar{I}$ and $\bar{J}$ denote the complements of I and J, respectively. Notice that $x_i F_i(x) = min\{x_i,\ F_i(x)\} \cdot max\{x_i,\ F_i(x)\}$. Consider the following four cases:

Case 1: $i \in \bar{I} \cap \bar{J}$. It follows that

$$\begin{aligned} & \alpha F_i(x) > x_i, \quad \alpha x_i > F_i(x). \\ \Rightarrow\quad & \alpha^2 F_i(x) > \alpha x_i > F_i(x). \\ \Rightarrow\quad & (\alpha^2 - 1)F_i(x) > 0. \\ \Rightarrow\quad & F_i(x) > 0,\ x_i > 0. \end{aligned}$$

We also have that

$$x_i F_i(x) \le \alpha min\{x_i^2, F_i(x)^2\}, \quad max\{x_i, F_i(x)\} \le \alpha min\{x_i,\ F_i(x)\}.$$

Therefore,

$$\begin{aligned}
M_i(x,\alpha) &= 2\alpha x_i F_i(x) - x_i^2 - F_i(x)^2 \\
&= 2\alpha x_i F_i(x) - 2x_i F_i(x) - (x_i - F_i(x))^2 \\
&\leq 2\alpha x_i F_i(x) - 2x_i F_i(x) \\
&= 2(\alpha - 1)x_i F_i(x) \\
&\leq 2\alpha(\alpha - 1)min\{x_i^2,\ F_i(x)^2\} \\
&= 2\alpha(\alpha - 1)(min\{x_i,\ F_i(x)\})^2 \\
&= 2\alpha(\alpha - 1)r_i(x)^2
\end{aligned}$$

On the other hand

$$\begin{aligned}
M_i(x,\alpha) &= 2\alpha x_i F_i(x) - x_i^2 - F_i(x)^2 \\
&= 2\alpha(x_i F_i(x) - (min\{x_i,\ F_i(x)\})^2) + 2\alpha(min\{x_i,\ F_i(x)\})^2 \\
&\quad -x_i^2 - F_i(x)^2 \\
&= 2\alpha min\{x_i,\ F_i(x)\}(max\{x_i,\ F_i(x)\} - min\{x_i,\ F_i(x)\}) \\
&\quad -x_i^2 - F_i(x)^2 + 2\alpha(min\{x_i,\ F_i(x)\})^2 \\
&\geq 2max\{x_i,\ F_i(x)\}(max\{x_i,\ F_i(x)\} - min\{x_i,\ F_i(x)\}) \\
&\quad -x_i^2 - F_i(x)^2 + 2\alpha r_i(x)^2 \\
&= 2(max\{x_i,\ F_i(x)\})^2 - 2x_i F_i(x) - x_i^2 - F_i(x)^2 + 2\alpha r_i(x)^2 \\
&= (max\{x_i,\ F_i(x)\})^2 - 2x_i F_i(x) - (min\{x_i, F_i(x)\})^2 \\
&\quad +2\alpha r_i(x)^2 \\
&= (x_i - F_i(x))^2 - 2(min\{x_i,\ F_i(x)\})^2 + 2\alpha r_i(x)^2 \\
&\geq -2(min\{x_i,\ F_i(x)\})^2 + 2\alpha r_i(x)^2 \\
&= 2(\alpha - 1)r_i(x)^2.
\end{aligned}$$

Case 2: $i \in I \cap J$. It follows that

$$\begin{aligned} & x_i \geq \alpha F_i(x), \quad F_i(x) \geq \alpha x_i. \\ \Rightarrow \quad & x_i \geq \alpha F_i(x) \geq \alpha^2 x_i. \\ \Rightarrow \quad & (1-\alpha^2)x_i \geq 0. \\ \Rightarrow \quad & x_i \leq 0, \ F_i(x) \leq 0. \end{aligned}$$

We also have that

$$\alpha x_i F_i(x) \geq (min\{x_i,\ F_i(x)\})^2, \quad min\{x_i,\ F_i(x)\} \geq \alpha max\{x_i,\ F_i(x)\}.$$

Therefore

$$\begin{aligned} M_i(x,\alpha) &= -2\alpha x_i F_i(x) + (\alpha x_i)^2 + (\alpha F_i(x))^2 \\ &= -2\alpha(x_i F_i(x) - (min\{x_i,\ F_i(x)\})^2) - 2\alpha(min\{x_i,\ F_i(x)\})^2 \\ &\quad +(\alpha x_i)^2 + (\alpha F_i(x))^2 \\ &= -2\alpha min\{x_i,\ F_i(x)\}(max\{x_i,\ F_i(x)\} - min\{x_i,\ F_i(x)\}) \\ &\quad -2\alpha r_i(x)^2 + (\alpha x_i)^2 + (\alpha F_i(x))^2 \\ &\leq -2\alpha^2 max\{x_i,\ F_i(x)\}(max\{x_i,\ F_i(x)\} - min\{x_i,\ F_i(x)\}) \\ &\quad -2\alpha r_i(x)^2 + (\alpha x_i)^2 + (\alpha F_i(x))^2 \\ &= -2(\alpha max\{x_i,\ F_i(x)\})^2 + 2\alpha^2 x_i F_i(x) - 2\alpha r_i(x)^2 \\ &\quad +(\alpha x_i)^2 + (\alpha F_i(x))^2 \\ &= -(\alpha max\{x_i,\ F_i(x)\})^2 + 2\alpha^2 x_i F_i(x) - 2\alpha r_i(x)^2 \\ &\quad +(\alpha min\{x_i,\ F_i(x)\})^2 \\ &= -\alpha^2(x_i - F_i(x))^2 + 2(\alpha min\{x_i(x),\ F_i(x)\})^2 - 2\alpha r_i(x)^2 \\ &\leq 2(\alpha r_i(x))^2 - 2\alpha r_i(x)^2 \\ &= 2\alpha(\alpha - 1) r_i(x)^2 \end{aligned}$$

On the other hand

$$\begin{aligned} M_i(x,\alpha) &= -2\alpha x_i F_i(x) + (\alpha x_i)^2 + (\alpha F_i(x))^2 \\ &= -2\alpha x_i F_i(x) + 2\alpha^2 x_i F_i(x) + \alpha^2 (x_i - F_i(x))^2 \\ &\geq 2(\alpha - 1)\alpha x_i F_i(x) \\ &\geq 2(\alpha - 1)(min\{x_i,\ F_i(x)\})^2 \\ &= 2(\alpha - 1) r_i(x)^2. \end{aligned}$$

Case3: $i \in I \cap \bar{J}$. Then it follows that

$$\begin{aligned} & x_i \geq \alpha F_i(x),\ \alpha x_i > F_i(x). \\ \Rightarrow\ & x_i + \alpha x_i > \alpha F_i(x) + F_i(x). \\ \Rightarrow\ & x_i > F_i(x). \\ \Rightarrow\ & r_i(x) = F_i(x). \end{aligned}$$

Therefore

$$\begin{aligned} M_i(x,\alpha) &= (\alpha F_i(x))^2 - F_i(x)^2 \\ &= (\alpha^2 - 1) r_i(x)^2 \\ &\leq 2\alpha(\alpha - 1) r_i(x)^2. \end{aligned}$$

On the other hand

$$\begin{aligned} M_i(x,\alpha) &= (\alpha^2 - 1) r_i(x)^2 \\ &\geq 2(\alpha - 1) r_i(x)^2. \end{aligned}$$

Case 4: $i \in \bar{I} \cap J$. It follows that

$$\alpha F_i(x) > x_i,\ F_i(x) \geq \alpha x_i.$$

$$\Rightarrow \quad \alpha F_i(x) + F_i(x) > x_i + \alpha x_i.$$

$$\Rightarrow \quad F_i(x) > x_i.$$

$$\Rightarrow \quad r_i(x) = x_i.$$

Therefore

$$\begin{aligned} M_i(x,\alpha) &= (\alpha x_i)^2 - x_i^2 \\ &= (\alpha^2 - 1) r_i(x)^2 \\ &\leq 2\alpha(\alpha - 1) r_i(x)^2. \end{aligned}$$

On the other hand

$$\begin{aligned} M_i(x,\alpha) &= (\alpha^2 - 1) r_i(x)^2 \\ &\geq 2(\alpha - 1) r_i(x)^2. \end{aligned}$$

From the four cases above we conclude that

$$2(\alpha-1)\sum_{i=1}^{n} r_i(x)^2 \leq \sum_{i=1}^{n} M_i(x,\alpha) \leq 2\alpha(\alpha-1)\sum_{i=1}^{n} r_i(x)^2,$$

that is

$$2(\alpha-1) r(x)^2 \leq M(x,\alpha) \leq 2\alpha(\alpha-1) r(x)^2.$$

Q.E.D.

From the theorem above and Theorem 2.2.2, the following theorem follows immediately.

Theorem 4.2.2 *Let $F(x) = Mx + q$ for $M \in \mathbb{R}^{n\times n}$. For each $\alpha > 1$, there exists some constants $\kappa > 0$ and $\delta > 0$ such that*

$$\|x - \bar{x}(x)\| \leq \kappa (M(x,\alpha))^{\frac{1}{2}}, \quad \text{whenever } M(x,\ \alpha) \leq \delta, \tag{4.2.2}$$

where $\bar{x}(x)$ is an orthogonal projection of x on the solution set of the LCP (2.1.1).

Proof. The proof follows directly from Theorem 2.2.2 and 4.2.1. **Q.E.D.**

We now show that if $F(x)$ is strongly monotone, then $(M(x,\alpha))^{\frac{1}{2}}$ is a global error bound for the NCP (3.1.1).

Theorem 4.2.3 *Let $F(x)$ be strongly monotone and Lipschitz continuous. Then for each $\alpha > 1$, there exists a constant $\tau > 0$ such that*

$$\|x - \bar{x}(x)\| \leq \tau(M(x,\alpha))^{\frac{1}{2}}, \tag{4.2.3}$$

where $\bar{x}$ is the unique solution of the NCP (3.1.1).

Proof. The proof is easy to see by Theorem 3.2.1 and 4.2.1. **Q.E.D.**

Finally Luo [LMRS92] indicated that $(M'(x,\alpha))^{\frac{1}{2}}$, where the prime indicates differentiation with respect to α, is also a local error bound. In fact by a proof similar to that of Theorem 4.2.1, we can further establish the following interesting relation between $M'(x,\alpha)$ and $r(x)$

$$2r(x)^2 \leq M'(x,\alpha) \leq 2\alpha r(x)^2$$

.

4.3 Relating the Restricted Implicit Lagrangian to Other Error Residuals

When restricted to the nonnegative orthant, the implicit Lagrangian becomes the restricted implicit Lagrangian ([MaS92], [Fuk92]) defined as follows

$$N(x,\alpha) = 2\alpha x F(x) + \|(-\alpha F(x) + x)_+\|^2 - \|x\|^2. \tag{4.3.1}$$

In this section we establish relationships between $N(x,\alpha)$ and known error residuals over $I\!R^n_+$. As a result, we obtain error bound results for $N(x,\alpha)$.

Theorem 4.3.1 *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be any function. Then for each $\alpha > 1$*

$$\alpha r(x)^2 \le N(x, \alpha), \quad \textit{for } x \ge 0. \tag{4.3.2}$$

Proof. Let $I = \{i \mid x_i \ge F_i(x)\}$ and

$$N_i(x, \alpha) = 2\alpha x_i F_i(x) + (-\alpha F_i(x) + x_i)_+^2 - x_i^2$$
$$r_i(x) = x_i - (x_i - F_i(x))_+ = \min\{x_i, F_i(x)\}.$$

For $i = 1, \ldots, n$, If $i \in I, x_i \ge \alpha F_i(x)$. Since $x_i \ge 0$, it is easy to see that $r_i(x) = F_i(x)$. Therefore,

$$\begin{aligned} N_i(x, \alpha) &= 2\alpha x_i F_i(x) + (-\alpha F_i(x) + x_i)^2 - x_i^2 \\ &= 2\alpha x_i F_i(x) + (\alpha F_i(x))^2 - 2\alpha x_i F_i(x) + x_i^2 - x_i^2 \\ &= (\alpha F_i(x))^2 \\ &= \alpha (F_i(x))^2 \\ &= \alpha r_i(x)^2. \end{aligned}$$

If $i \notin I, \alpha F_i(x) > x_i \ge 0$. Therefore

$$\begin{aligned} N_i(x, \alpha) &= 2\alpha x_i F_i(x) - x_i^2 \\ &\ge 2\alpha x_i F_i(x) - \alpha x_i F_i(x) \\ &= \alpha x_i F_i(x) \\ &\ge \alpha (\min\{x_i, F_i(x)\})^2 \\ &= \alpha r_i(x)^2. \end{aligned}$$

Hence, for $i = 1, \ldots, n$

$$\alpha r_i(x)^2 \le N_i(x, \alpha), \quad \text{for } x \ge 0.$$

By summing both sides on i from 1 to n, (4.3.2) holds. **Q.E.D.**

The following error bound results involving $N(x,\alpha)$ are immediate consequences of Theorem 4.3.1.

Corollary 4.3.1 *Let $F(x) = Mx+q$ for $M \in I\!R^{n\times n}$. For each $\alpha > 1$, $(N(x,\alpha))^{\frac{1}{2}}$ is a local error bound for $LCP(M,q)$ on the nonnegative orthant, i.e. there exist positive ϵ and τ such that*

$$\|x - \bar{x}(x)\| \leq \tau(N(x,\alpha))^{\frac{1}{2}}, \quad \text{for } N(x,\alpha) \leq \epsilon,\ x \geq 0,$$

where $\bar{x}(x)$ is a closest solution to x under norm $\|\cdot\|$.

Proof. The proof follows directly from Theorem 4.3.1 and Lemma 2.2.2. **Q.E.D.**

Corollary 4.3.2 *Let $F(x) = Mx+q$ for $M \in R_0$. For each $\alpha > 1$, $(N(x,\alpha))^{\frac{1}{2}}$ is a global error bound for $LCP(M,q)$ on the nonnegative orthant, i.e. there exists positive τ*

$$\|x - \bar{x}(x)\| \leq \tau(N(x,\alpha))^{\frac{1}{2}}, \quad \text{for } x \geq 0,$$

where $\bar{x}(x)$ is a closest solution to x under norm $\|\cdot\|$.

Proof. The proof follows directly from Theorem 4.3.1 and Theorem 2.2.1. **Q.E.D.**

Corollary 4.3.3 *Let $F(x)$ be strongly monotone and Lipschitz continuous. Assume that $\bar{x}$ is the unique solution for NCP. Then $(N(x,\alpha))^{\frac{1}{2}}$ is a global error bound on the nonnegative orthant, i.e. there exists positive λ such that*

$$\|x - \bar{x}\| \leq \lambda(N(x,\alpha))^{\frac{1}{2}}, \quad x \geq 0.$$

Proof. The proof follows directly from Theorem 4.3.1 and Theorem 3.2.1. **Q.E.D.**

On the other hand, $N(x, \alpha)$ can be bounded by error residuals we introduced before.

Theorem 4.3.2 *Let $F : I\!R^n \to I\!R^n$ be any function. Then for each $\alpha > 1$*

$$N(x, \alpha) \leq \alpha^2 r(x)^2 + 2\alpha \sum_{i=1}^{n} (x_i F_i(x))_+, \quad \text{for } x \geq 0. \tag{4.3.3}$$

Hence,

$$N(x, \alpha) \leq \alpha^2 r(x)^2 + 2\alpha t(x).$$

Proof. Let $I, r_i(x)$ and $N_i(x, \alpha)$ be defined in the same way as in Theorem 4.3.1. If $i \in I$, it is easy to see that $r_i(x) = F_i(x)$ and

$$N_i(x, \alpha) = \alpha^2 (F_i(x))^2 = \alpha^2 r_i(x)^2, \quad x \geq 0.$$

If $i \notin I$ and $x \geq 0$, then

$$N_i(x, \alpha) = 2\alpha x_i F_i(x) - x_i^2 \leq 2\alpha (x_i F_i(x))_+.$$

Hence,

$$N_i(x, \alpha) \leq \alpha^2 r_i(x)^2 + 2\alpha (x_i F_i(x))_+, \quad x \geq 0.$$

By summing up two sides above with respect to index i, then we have (4.3.3). **Q.E.D.**

Chapter 5

Error Bounds Generated by Various Algorithms

5.1 Introduction

In this chapter, we apply the error bound results obtained in Chapters 2 and 3 to approximate solutions generated by algorithms which are often used to solve optimization problems. The algorithms under consideration are exterior penalty, interior penalty, augmented Lagrangian and proximal point. As a result, we give a bound on the distance between each approximate algorithm-generated solution and a closest real solution in terms of computable quantities. In some instances, the bound is in terms of the penalty parameter of a penalty function. In others the bound depends on a measure of infeasibility and complementarity. These bounds are easily available. They can serve as a termination criteria for various algorithms. They also can be used as a guide for improvement of an approximate solution since these bounds explicitly give the quantities that affect the accuracy of the approximate solution.

In Section 5.2, methods such as penalty function or augmented Lagrangian are used to solve a linear program. We obtain a linear relation between the distance from an algorithm-generated point to the solution set and quantities computed by various methods. In Section 5.3, we use a similar approach to solve a convex quadratic program. Similar bounds on the actual distance are obtained except that in some cases, the linear dependence relation no longer holds and additional terms are necessary to bound the distance. In Sections 5.4 and 5.5, the same approach is applied to a more complicated strongly monotone nonlinear complementarity problem and a strongly convex program, respectively. We extract the same type of algorithm-generated error bound results.

5.2 Error Bounds for Linear Programs

In this section, various computational approaches are applied to a linear program. Actual error bounds between any inexact solution of each approach and the solution set of the linear program are obtained. The new idea here is to use computable quantities obtained from these approaches to linearly bound the distance to the solution set.

Exterior Penalty Approach Consider the linear program:

$$\begin{aligned} min \quad & cx \\ s.t. \quad & Ax \leq b \end{aligned} \tag{5.2.1}$$

and its dual

$$\begin{aligned} max \quad & -bu \\ s.t. \quad & A^T u = -c \end{aligned}$$

$$u \geq 0. \tag{5.2.2}$$

Define the associated exterior penalty function for (5.2.1) as follows

$$P(x, \alpha) := cx + \frac{\alpha}{2}\|(Ax - b)_+\|_2^2, \quad \alpha > 0. \tag{5.2.3}$$

Let $x(\alpha)$ be a minimizer of this exterior function. From the theory of penalty functions ([Ma86] [FiM68]), $x(\alpha)$ asymptotically goes to the solution set of the linear program (5.2.1). However, these results do not relate $x(\alpha)$ explicitly to a solution of the problem for a given α. In this thesis we bound the distance between $x(\alpha)$ and a closest solution of the linear program (5.2.1). The bounds we will develop answer the above question. In order to develop such a bound, first we cite the famous Hoffman theorem ([Hof52], also [Rob73] [MaS87]) which gives a bound on the distance between any $x \in I\!R^n$ and a closest solution of the linear program (5.2.1).

Lemma 5.2.1 *([Hof52] [Rob81] [MaS87]) Let the solution set of (5.2.1) be non-empty. Then, there exists a* $\sigma(A, b, c) > 0$ *such that* $\forall (x, u) \in I\!R^{n+m}$

$$\|(x, u) - (\bar{x}(x, u), \bar{u}(x, u))\| \leq \sigma(A, b, c)\|(cx + bu,\ Ax - b,\ -u)_+,\ A^T u + c\|, \tag{5.2.4}$$

where $(\bar{x}(x, u), \bar{u}(x, u))$ *is a closest KKT pair of the linear program (5.2.1) from* (x, u) *under the norm* $\|\cdot\|$.

The following lemma gives a very interesting result that the right-hand side norm in (5.2.4) also provides a lower bound for the distance from (x, u) to the KKT pair set of the linear program (5.2.1). In other words, the distance is equivalent to the right-hand norm: that is, they bound each other. This tells that this computable norm quantity well characterizes the distance and can be easily calculated.

Lemma 5.2.2 *Let the solution set of (5.2.1) be nonempty and the norm used here is monotonic. Then, there exists a* $\mu > 0$ *such that* $\forall (x,u) \in I\!R^{n+m}$

$$\|(x,u) - (\bar{x},\bar{u})\| \geq \mu \|(cx + bu,\ Ax - b,\ -u)_+,\ A^T u + c\|, \tag{5.2.5}$$

where $(\bar{x},\bar{u})$ *is any KKT pair of the linear program (5.2.1).*

Proof. Since $c\bar{x} + b\bar{u} = 0, A\bar{x} \leq b, u \geq 0$ and $A^T\bar{u} + c = 0$, so

$$\begin{aligned}
&\|((cx + bu,\ Ax - b,\ -u)_+,\ A^T u + c)\| \\
&= \|((cx + bu,\ Ax - b,\ -u)_+,\ A^T u + c) - ((c\bar{x} + b\bar{u},\ A\bar{x} - b, \bar{u})_+,\ A^T\bar{u} + c)\| \\
&\leq \|c(x - \bar{x}) + b(u - \bar{u}),\ A(x - \bar{x}),\ -(u - \bar{u}),\ A^T(u - \bar{u})\| \\
&\leq \tau(A,b,c)\|(x,u) - (\bar{x},\bar{u})\|,
\end{aligned}$$

where $\tau(A,b,c)$ is some positive constant. Take $\mu = \frac{1}{\tau(A,b,c)}$, then (5.2.5) follows. **Q.E.D.**

Before deriving a bound on the distance between $x(\alpha)$ and the solution set of the linear program (5.2.1), we need to establish the existence of such an $x(\alpha)$ with the following lemma.

Lemma 5.2.3 *Let the linear program (5.2.1) be solvable. Then the penalty function (5.2.3) has a minimizer* $x(\alpha)$ *for each* $\alpha > 0$.

Proof. Minimizing $P(x,\alpha)$ on $I\!R^n$ is equivalent to ([Ma83])

$$\begin{aligned}
min \quad & P(x,y,\alpha) := cx + \tfrac{\alpha}{2}\|Ax - b + y\|_2^2 \\
s.t. \quad & y \geq 0.
\end{aligned} \tag{5.2.6}$$

It is easy to see that (5.2.6) is a convex quadratic program. If it is both primal and dual feasible, then it is solvable ([FraWol56]). Obviously (5.2.6) is primal feasible. So the only thing we need prove is to find a dual feasible point. Consider the dual quadratic program

$$\begin{aligned} &min && P(x,y,\alpha) - yu \\ &s.t.\ \nabla_{(x,y)}(P(x,y,\alpha) - yu) = && \begin{bmatrix} c + \alpha A^T(Ax - b + y) \\ \alpha(Ax - b + y) - u \end{bmatrix} = 0 \\ & && u \geq 0. \end{aligned} \tag{5.2.7}$$

Since (5.2.1) is solvable, there exist $\bar{x}$ and $\bar{u}$ that solve (5.2.1) and (5.2.2) respectively. Now in (5.2.7), take $x = \bar{x}$, $u = \bar{u} \geq 0$ and $y = \bar{u}/\alpha - A\bar{x} + b$, then the constraints in (5.2.7) are satisfied. In other words, (5.2.6) is dual feasible. Hence the primal objective is bounded below and (5.2.3) has a minimum solution point for $\alpha > 0$. **Q.E.D.**

The following theorem estimates for a fixed value of the penalty parameter $\alpha > 0$, the distance between a minimum solution to the exterior penalty function (5.2.3) and an exact solution to the linear program (5.2.1).

Theorem 5.2.1 *Suppose the linear program (5.2.1) is solvable. Let $x(\alpha)$ be a minimizer of $P(x,\alpha)$. Then there exists a constant $\sigma(A,b,c) > 0$ such that*

$$\|x(\alpha) - \bar{x}(x(\alpha))\| \leq \sigma(A,b,c)\|(Ax(\alpha) - b)_+\|, \tag{5.2.8}$$

where $\bar{x}(x(\alpha))$ is the orthogonal projection of $x(\alpha)$ on the solution set of (5.2.1).

Proof. By Lemma 5.2.3, such a $x(\alpha)$ exists. It satisfies the first order optimality condition

$$\nabla P(x(\alpha),\alpha) = c + \alpha A^T(Ax(\alpha) - b)_+ = 0.$$

Let $u = \alpha(Ax(\alpha) - b)_+$, then it follows

$$(cx(\alpha) + bu)_+ \quad = \quad (cx(\alpha) + x(\alpha)A^T u - x(\alpha)A^T u + bu)_+$$

$$
\begin{aligned}
&= (x(\alpha)^T(c + A^T u) + (b - Ax(\alpha))^T u)_+ \\
&= ((b - Ax(\alpha))^T u)_+ \\
&= \alpha((b - Ax(\alpha))^T (Ax(\alpha) - b)_+)_+ \\
&= 0.
\end{aligned}
$$

This together with Lemma 5.2.1 leads to

$$
\begin{aligned}
\|x(\alpha) - \bar{x}(x(\alpha))\| &\leq \sigma(A, b, c)\|(cx(\alpha) + bu,\ Ax(\alpha) - b,\ -u)_+,\ A^T u + c\| \\
&= \sigma(A, b, c)\|(Ax(\alpha) - b)_+\|.
\end{aligned}
$$

Q.E.D.

Theorem 5.2.1 shows that the error in $x(\alpha)$ as measured by its distance from the solution set of the linear program, is linearly dependent on the violation of constraints of the linear program. Thus, if we can cut the violation by one half, then the distance is automatically reduced by one half.

Interior Penalty Approach Consider the linear program

$$
\begin{aligned}
\min \quad & cx \\
s.t. \quad & Ax = b \\
& x \geq 0
\end{aligned}
\tag{5.2.9}
$$

and its dual

$$
\begin{aligned}
\max \quad & bu \\
s.t. \quad & A^T u \leq c.
\end{aligned}
\tag{5.2.10}
$$

Define the associated interior penalty minimization problem with (5.2.9) as follows

$$
\min \quad Q(x, \gamma) := cx - \gamma \textstyle\sum_{i=1}^{n} \log x_i
$$

$$s.t. \qquad Ax = b$$
$$x > 0. \tag{5.2.11}$$

Let $x(\gamma)$ be a solution of (5.2.11). In order to bound the distance estimate between $x(\gamma)$ and the solution set of the linear program (5.2.9), we first need the following lemma that is similar to Lemma 5.2.1.

Lemma 5.2.4 *([Hof52] [Rob73] [MaS87]) Let (5.2.9) be solvable. Then, there exists a $\sigma(A, b, c) > 0$ such that $\forall (x, u) \in \mathbb{R}^{n+m}$*

$$\|x - \bar{x}(x)\| \le \sigma(A, b, c)\|(cx - bu,\ A^T u - c,\ -x)_+,\ Ax - b\|,$$

where $\bar{x}(x)$ is the orthogonal projection of x on the solution set of (5.2.9).

Proof. The proof is similar to that of Lemma 5.2.1. **Q.E.D.**

Now by using the above lemma, we obtain the following interesting bound in terms of the interior penalty parameter γ. In particular this parameter itself is enough to bound the distance between $x(\gamma)$,solution to (5.2.11) and the solution set of the linear program (5.2.9). Note that, unlike the exterior case, the solution $x(\gamma)$ of (5.2.11) may not exist even if the linear program (5.2.9) is solvable. However, (5.2.11) is solvable under a variety of sufficient conditions such as boundedness of the primal feasible region.

Theorem 5.2.2 *Let both (5.2.9) and (5.2.11) be solvable. Suppose $x(\gamma)$ is a solution of (5.2.11). Then, there exists a $\sigma(A, b, c) > 0$ such that*

$$\|x(\gamma) - \bar{x}(x(\gamma))\| \le \sigma(A, b, c)\gamma n, \tag{5.2.12}$$

where $\bar{x}(x(\gamma))$ is the orthogonal projection of $x(\gamma)$ on the solution set of (5.2.9).

Proof. Since $x(\gamma)$ is an optimal solution, there is a $u \in I\!R^m$ such that

$$c - \gamma X^{-1}e - A^T u = 0,$$

where $X = diag(x(\gamma))$, $x(\gamma) > 0$ and $e = (1, 1, \ldots, 1)$. By Lemma 5.2.4, it follows that

$$\begin{aligned}
\|x(\gamma) - \bar{x}(x(\gamma))\| &\leq \sigma(A, b, c)\|(cx(\gamma) - bu,\ A^T u - c,\ -x(\gamma))_+, \\
&\quad Ax(\gamma) - b\| \\
&= \sigma(A, b, c)(cx(\gamma) - bu)_+ \\
&= \sigma(A, b, c)(cx(\gamma) - \gamma x(\gamma)^T X^{-1}e - b^T u \\
&\quad +\gamma x(\gamma)^T X^{-1} e)_+ \\
&= \sigma(A, b, c)(cx(\gamma) - \gamma x(\gamma)^T X^{-1}e - x(\gamma) A^T u \\
&\quad +\gamma x(\gamma)^T X^{-1} e)_+ \\
&= \sigma(A, b, c)(\gamma x(\gamma)^T X^{-1} e)_+ \\
&= \sigma(A, b, c)\gamma n.
\end{aligned}$$

Q.E.D.

The theorem also justifies the use of penalty function methods ([FiM68], [BaS79]) by showing that distance to the solution set from $x(\gamma)$ is bounded linearly by γ.

Augmented Lagrangian Approach Consider the linear program (5.2.1) and define the associated augmented Lagrangian as follows

$$L(x, u, \alpha) := cx + \frac{1}{2\alpha}[\|(\alpha(Ax - b) + u)_+\|_2^2 - \|u\|_2^2], \quad \alpha > 0. \tag{5.2.13}$$

Let $x(\alpha, u)$ be a minimizer of (5.2.13) with respect to x for fixed $u \geq 0$ and fixed $\alpha > 0$. First, we establish the existence of such an $x(\alpha, u)$.

Lemma 5.2.5 *Let (5.2.1) be solvable. Then (5.2.13) has a minimizer with respect to x for each $\alpha > 0$ and $u \geq 0 \in I\!R^m$.*

Proof. The is similar to that of lemma 5.2.3. **Q.E.D.**

From augmented Lagrangian theory [Hes69] [Pow69], the unconstrained minimizer $x(u, \alpha)$ with respect to x of (5.2.13) is a true solution of the linear program (5.2.1) when an exact optimal multiplier is used for u for $\alpha > 0$. Often $x(u, \alpha)$ is used as an approximate solution of (5.2.1) in practice without knowing an exact multiplier. However there was previously no good way to know the accuracy of such $x(u, \alpha)$. The following theorem provides a computable bound on the distance between $x(u, \alpha)$ and the solution set of (5.2.1). Moreover it explicitly gives the quantities that affect the error.

Theorem 5.2.3 *Let (5.2.1) be solvable and let $x(\alpha, u)$ be a minimizer of the augmented Lagrangian (5.2.13) respect to x for some $u \geq 0$ and $\alpha > 0$. Assume that the following norm is monotonic. Then there exists $\sigma(A, b, c) > 0$ such that*

$$\begin{aligned} \|x(\alpha, u) - \bar{x}(x(\alpha, u))\| \leq \; & \sigma(A, b, c)\|(b - Ax(\alpha, u))_+ (\alpha(Ax(\alpha, u) - b) + u)_+, \\ & (Ax(\alpha, u) - b)_+\|, \end{aligned} \tag{5.2.14}$$

where $\bar{x}(x(\alpha, u))$ is the orthogonal projection of $x(\alpha, u)$ on the solution set of (5.2.1).

Proof. By Lemma 5.2.5, such a minimizer $x(\alpha, u)$ exists. It satisfies the first order optimality condition

$$\nabla_x L(x(\alpha, u), u, \alpha) = c + A^T[\alpha(Ax(\alpha, u) - b) + u]_+ = 0. \tag{5.2.15}$$

Hence the point $[\alpha(Ax(\alpha,u)-b)+u]_+$ is dual feasible for the linear program (5.2.1). Therefore,

$$\begin{aligned}
& cx(\alpha,u)+b(\alpha(Ax(\alpha,u)-b)+u)_+ \\
&= cx(\alpha,u)+x(\alpha,u)A^T[\alpha(Ax(\alpha,u)-b)+u]_+ \\
&\quad -x(\alpha,u)A^T[\alpha(Ax(\alpha,u)-b)+u]_+ + b(\alpha(Ax(\alpha,u)-b)+u)_+ \\
&= (b-Ax(\alpha,u))[\alpha(Ax(\alpha,u)-b)+u]_+.
\end{aligned}$$

By Lemma 5.2.1, we have

$$\begin{aligned}
& \|x(\alpha,u)-\bar{x}(x(\alpha,u))\| \leq \\
& \quad \sigma(A,b,c)\|cx(\alpha,u)+b(\alpha(Ax(\alpha,u)-b)+u)_+,(Ax(\alpha,u)-b)_+\| \\
& \leq \sigma(A,b,c)\|(b-Ax(\alpha,u))_+[\alpha(Ax(\alpha,u)-b)+u]_+,(Ax(\alpha,u)-b)_+\|.
\end{aligned}$$

Q.E.D.

Theorem 5.2.3 also illustrates that if $x(\alpha,u)$ is feasible for the linear program (5.2.1) and

$$(b-Ax(\alpha,u))_+(\alpha(Ax(\alpha,u)-b)+u)_+=0$$

then $x(\alpha,u)$ is an exact solution of the linear program. This justifies the result of augmented Lagrangian theory, that if $x(\alpha,u)$ is primal feasible and

$$u=(\alpha(Ax(\alpha,u)-b)+u)_+, \quad u(b-Ax(\alpha,u))=0. \tag{5.2.16}$$

then $x(\alpha,u)$ is an exact solution of the linear program (5.2.1).

Proximal Point Approach Consider the linear program (5.2.1) and define the associated proximal point minimization problem as follows

$$\begin{aligned}
&\min \quad P(x,y,\epsilon):=cx+\frac{\epsilon}{2}\|x-y\|_2^2, \quad \epsilon>0, \quad \text{for fixed y} \\
&s.t. \quad Ax\leq b.
\end{aligned} \tag{5.2.17}$$

Note that this problem is always solvable since the objective function is strongly convex. Now we obtain the following theorem that gives a computable bound on the distance between the unique solution of (5.2.17) and the solution set of the linear program (5.2.1).

Theorem 5.2.4 *Let (5.2.1) be solvable and $x(\epsilon, y)$ be a solution of (5.2.17) under the constraints for some $\epsilon > 0$ and $y \in I\!R^n$. Then there exists a $\sigma(A, b, c) > 0$ such that*

$$\|x(\epsilon, y) - \bar{x}(x(\epsilon, y))\| \leq \sigma(A, b, c)\|(-\epsilon x(\epsilon, y)(x(\epsilon, y) - y))_+, -\epsilon(x(\epsilon, y) - y)\| \tag{5.2.18}$$

where $\bar{x}(x(\epsilon, y))$ is the orthogonal projection of $x(\epsilon, y)$ on the solution set of (5.2.1).

Proof. Since $x(\epsilon, y)$ solves (5.2.17), there exits some $u \geq 0 \in I\!R^m$ such that

$$\begin{aligned} \nabla_x P(x(\epsilon, y)), y, \epsilon) + A^T u &= c + \epsilon(x(\epsilon, y)) - y) + A^T u = 0 \\ & Ax(\epsilon, y) \leq b, \quad u \geq 0 \\ & u(Ax(\epsilon, y) - b) = 0. \end{aligned}$$

By Lemma 5.2.1, there exists a $\sigma(A, b, c) > 0$ such that

$$\begin{aligned} \|x(\epsilon, y) - \bar{x}(x(\epsilon, y))\| &\leq \sigma(A, b, c)\|(cx(\epsilon, y) + bu,\ Ax(\epsilon, y) - b,\ -u)_+, \\ & \quad A^T u + c\| \\ &= \sigma(A, b, c)\|(cx(\epsilon, y) + bu)_+,\ -\epsilon(x(\epsilon, y) - y)\|. \end{aligned}$$

The term $(cx(\epsilon, y) + bu)_+$ can be further simplified as follows

$$\begin{aligned} (cx(\epsilon, y) + bu)_+ &= (cx(\epsilon, y) + \epsilon x(\epsilon, y)(x(\epsilon, y) - y) + bu - \epsilon x(\epsilon, y)(x(\epsilon, y) - y))_+ \\ &= (-\epsilon x(\epsilon, y)(x(\epsilon, y) - y))_+. \end{aligned}$$

Hence we have the final inequality

$$\|x(\epsilon,y)-\bar{x}(x(\epsilon,y))\| \leq \sigma(A,b,c)\|(-\epsilon x(\epsilon,y)(x(\epsilon,y)-y))_+, -\epsilon(x(\epsilon,y)-y)\|.$$

Q.E.D.

This theorem shows that the distance between the unique solution of the problem (5.2.17) and the solution set the linear program depends on the size of the ϵ and the closeness between $x(\epsilon, y)$ and y. Unfortunately, it does not include the result that when ϵ is small enough, then the proximal solution $x(\epsilon, y)$ is actually an exact solution of the linear program [MaM79]. **Q.E.D.**

5.3 Error Bounds for Quadratic Programs

In this section we bound the distance between each algorithm-generated approximate solution and the solution set of a convex quadratic program. The bounds we get, exhibit dependence relations on computable constraint violations, known penalty parameters and known quantities.

The way we obtain these bounds is by adapting the error bounds obtained earlier, such as $r(x) + s(x)$ and $s(x) + s(x)^{\frac{1}{2}}$ discussed in Chapter 2 to optimality conditions of convex quadratic programs. Note that Hoffman's theorem does not apply in this case. Consequently, the error bounds here may contain additional nonlinear terms of penalty parameters and violations of the constraints, that were not present in the case of the linear programs discussed in the last section.

Consider the quadratic program

$$\begin{aligned} &min \quad \tfrac{1}{2}xQx + cx \\ &s.t. \qquad Ax \geq b. \end{aligned} \tag{5.3.1}$$

where $Q \in I\!R^{n\times n}$ is symmetric positive semi-definite (spsd for short), $A \in I\!R^{m\times n}$, $c \in I\!R^n$ and $b \in I\!R^m$. Assume that (5.3.1) has a nonempty solution set.

We begin by giving an error bound in terms of the violations of the KKT conditions for (5.3.1) by using $s(x) + s(x)^{\frac{1}{2}}$.

Lemma 5.3.1 *There exists a $\sigma(Q,A,b,c) > 0$ such that for all $(x,u) \in I\!R^{n+m}$,*

$$\begin{aligned}&\|(x,u) - (\bar{x}(x,u), \bar{u}(x,u))\| \leq \\ &\quad \sigma(Q,A,b,c)\|(-Ax+b,\ -u,\ x(Qx+c-A^Tu)+u^T(Ax-b))_+, \\ &\quad Qx+c-A^Tu\| + \|(-Ax+b,\ -u,\ x(Qx+c-A^Tu)+u^T(Ax-b))_+, \\ &\quad Qx+c-A^Tu\|^{\frac{1}{2}} \end{aligned} \tag{5.3.2}$$

where $(\bar{x}(x,u), \bar{u}(x,u))$ is the orthogonal projection of (x,u) on the set of the KKT pairs of (5.3.1).

Proof. Let $x = x_+ - x_-$ and $x_+,\ x_- \geq 0$, then (5.3.1) becomes

$$\begin{aligned} min \quad & \tfrac{1}{2}(x_+ - x_-)^T Q (x_+ - x_-) + c(x_+ - x_-) \\ s.t. \quad & A(x_+ - x_-) \geq b \\ & x_+,\ x_- \geq 0. \end{aligned}$$

This is equivalent to

$$\begin{aligned} min \quad & \tfrac{1}{2} z^T \bar{Q} z + \bar{c} z \\ s.t. \quad & \bar{A} z \geq b \\ & z \geq 0, \end{aligned} \tag{5.3.3}$$

where

$$\bar{Q} = \begin{bmatrix} Q & -Q \\ -Q & Q \end{bmatrix} .$$

and $\bar{A} = (A, -A)$, $z = (x_+, x_-)$, $\bar{c} = (c, -c)$. By the assumption that Q is spsd, it is easy to see that $\bar{Q}$ is spsd too. Therefore (5.3.3) is a convex quadratic program. Hence solving it with $\bar{z}$ is equivalent to satisfying the following first order optimality condition with $\bar{w} = (\bar{z}, \bar{u})$ for some $\bar{u} \in I\!R^m$

$$Mw + q \geq 0, \quad w \geq 0, \quad w(Mw+q) = 0, \tag{5.3.4}$$

where

$$M = \begin{bmatrix} \bar{Q} & -\bar{A}^T \\ \bar{A} & 0 \end{bmatrix}, \qquad q = \begin{bmatrix} \bar{c} \\ -b \end{bmatrix}.$$

This is an LCP with a matrix $M \in I\!R^{(2n+m)\times(2n+m)}$ and a vector $q \in I\!R^{2n+m}$. We can easily show that M is positive semidefinite. Therefore, by Theorem 2.3.1, there exits a $\sigma(Q, A, b, c)$ such that $\forall w \in I\!R^{2n+m}$

$$\|w - \bar{w}(w)\| \leq \sigma(Q, A, b, c)(s(w) + s(w)^{\frac{1}{2}}), \tag{5.3.5}$$

where $\bar{w}(w)$ is the orthogonal projection of w on the solution set of (5.3.4). Let $(x, u) \in I\!R^{n+m}$, then take $w = (x_+, x_-, u) \in I\!R^{2n+m}$ where $x_+ = (x)_+, x_- = (-x)_+$. For such a w there exists a $\bar{w}(w) = (\bar{x}_+(w), \bar{x}_-(w), \bar{u}(w))$ such that (5.3.5) holds. Since

$$s(w) = \|(-Mw - q, \ -w, \ w(Mw+q))_+\|$$

and

$$Mw + q = \begin{bmatrix} Qx + c - A^T u \\ -Qx - c + A^T u \\ Ax - b \end{bmatrix}, \qquad w = \begin{bmatrix} x_+ \\ x_- \\ u \end{bmatrix},$$

it follows that

$$s(w) = \|(-Ax + b, \ -u, \ x(Qx + c - A^T u) + u(Ax - b))_+, \ Qx + c - A^T u\|. \tag{5.3.6}$$

Now let $\bar{x}(w) = \bar{x}_+(w) - \bar{x}_-(w)$. By the definition at the begin of the proof, $(\bar{x}(w), \bar{u}(w))$ is a KKT pair of (5.3.1). It follows that

$$\begin{aligned} \|(x,u) - (\bar{x}(x,u), \bar{u}(x,u)\|_1 &\leq \|(x,u) - (\bar{x}(w), \bar{u}(w))\|_1 \\ &\leq \|w - \bar{w}(w)\|_1. \end{aligned}$$

By the equivalence of norms together with (5.3.5) and (5.3.6), we could conclude that (5.3.2) holds for any norm. **Q.E.D.**

The next lemma generates another error bound on the same distance by using $r(x) + s(x)$ instead of $s(x) + s(x)^{\frac{1}{2}}$ as discussed in Chapter 2. This will lead later on to different algorithm-generated error bounds that sometimes are better than that obtained by using $s(x) + s(x)^{\frac{1}{2}}$.

Lemma 5.3.2 *There exists a $\sigma(Q, A, b, c) > 0$ such that for all $(x,u) \in I\!\!R^{n+m}$,*

$$\|(x,u) - (\bar{x}(x,u), \bar{u}(x,u))\| \leq \sigma(Q, A, b, c)(r(w) + s(w)). \tag{5.3.7}$$

where $s(w)$ and $r(w)$ are defined in (5.3.6) and (5.3.8), respectively, and $(\bar{x}(x,u), \bar{u}(x,u))$ is the orthogonal projection of (x,u) on the set of the KKT pairs of (5.3.1).

Proof. In the proof of Lemma 5.3.1, let s(w) be the same as in (5.3.6), but

$$\begin{aligned} r(w) &= \|\min\{Mw + q,\ w\}\| \\ &= \|\min\{\begin{bmatrix} Qx + c - A^T u \\ -Qx - c + A^T u \\ Ax - b \end{bmatrix}, \begin{bmatrix} x_+ \\ x_- \\ u \end{bmatrix}\}\|. \end{aligned} \tag{5.3.8}$$

By Theorem 2.3.1, there exists a $\sigma(Q, A, b, c) > 0$ such that for all $(x,u) \in I\!\!R^{n+m}$,

$$\|(x,u) - (\bar{x}(x,u), \bar{u}(x,u))\| \leq \sigma(Q, A, b, c)(r(w) + s(w))$$

Q.E.D.

Exterior Penalty Function Consider the quadratic program (5.3.1) and the associated exterior penalty function

$$P(x,\alpha) := \frac{1}{2}xQx + cx + \frac{\alpha}{2}\|(-Ax+b)_+\|_2^2, \quad \alpha > 0. \tag{5.3.9}$$

First we show the existence of a minimizer of the function $P(x,\alpha)$ for $\alpha > 0$ in the following lemma.

Lemma 5.3.3 *Let the quadratic program (5.3.1) be solvable. Then $P(x,\alpha)$ has a minimizer.*

Proof. The proof is similar to that of Lemma 5.2.3. **Q.E.D.**

The new idea here is to relate the minimizer $x(\alpha)$ of (5.3.9) to the solution set of the quadratic program (5.3.1) by using Lemmas (5.3.1) and (5.3.2). This enables us to show a new error bound on the distance from each $x(\alpha)$ to the solution set of the quadratic program. In addition the error bound involves only the computable constraint violations. In the following theorem we derive these results.

Theorem 5.3.1 *Let the quadratic program (5.3.1) be solvable and $x(\alpha)$ be a minimizer of the penalty function (5.3.9), then there exists a $\sigma(Q,A,b,c) > 0$ such that*

$$\begin{aligned} \|x(\alpha) - \bar{x}(x(\alpha))\| \;\leq\; & \sigma(Q,A,c,b)(\|(-Ax(\alpha)+b)_+\| \\ & + \|(-Ax(\alpha)+b)_+\|^{1/2}) \end{aligned} \tag{5.3.10}$$

where $\bar{x}(x(\alpha))$ is the orthogonal projection of $x(\alpha)$ on the solution set of (5.3.1).

Proof. By Lemma 5.3.3, such a $x(\alpha)$ exists. It satisfies the first order optimality condition

$$\nabla_x P(x(\alpha),\alpha) = Qx + c - \alpha A^T(-Ax(\alpha)+b)_+ = 0.$$

By Lemma 5.3.1, (5.3.2) holds where $u = \alpha(-Ax(\alpha) + b)_+$. Since

$$(u(Ax(\alpha) - b))_+ = \alpha((-Ax(\alpha) + b)_+(Ax(\alpha) - b))_+ = 0$$

and

$$\|x(\alpha) - \bar{x}(x(\alpha))\| \leq \|(x(\alpha), u) - (\bar{x}(x(\alpha), u), \bar{u}(x(\alpha), u))\|,$$

where $(\bar{x}(x(\alpha), u), \bar{u}(x(\alpha), u))$ is defined in Lemma 5.3.1, we conclude that (5.3.10) holds. **Q.E.D.**

In (5.3.10) the error from $x(\alpha)$ to the solution set depends not only certain norm of the constraint violation, but also the additional square root of the same quantity. Therefore, if this violation size is small (< 1), the square root term actually increases the bound which is not desirable. Fortunately in this case, by using the Lemma 5.3.2, we derive in the following theorem a bound that depends linearly on the norm the constraint violation without involving a square root term that is the case of the linear program discussed in the last section.

Theorem 5.3.2 *Let the quadratic program (5.3.1) be solvable and $x(\alpha)$ be a minimizer of the penalty function (5.3.9). Assume that the following norm is monotonic. Then there exists a $\sigma(Q, A, b, c) > 0$ such that*

$$\|x(\alpha) - \bar{x}(x(\alpha))\| \leq 2\sigma(Q, A, c, b)\|(-Ax(\alpha) + b)_+\|. \tag{5.3.11}$$

where $\bar{x}(x(\alpha))$ is the orthogonal projection of $x(\alpha)$ on the solution set of (5.3.1).

Proof. Similar to the previous theorem, we have

$$s(w) = \|(-Ax(\alpha) + b)_+\|$$

By the definition (5.3.8) of $r(w)$ we further have

$$\begin{aligned} r(w) &= \| \min\{ \begin{bmatrix} Qx + c - A^T u \\ -Qx - c + A^T u \\ Ax - b \end{bmatrix}, \quad \begin{bmatrix} x_+ \\ x_- \\ u \end{bmatrix} \} \| \\ &= \| \min\{ Ax(\alpha) - b, \alpha(-Ax(\alpha) + b)_+ \} \| \\ &= \| (b - Ax(\alpha))_+ \|. \end{aligned}$$

Hence, by Lemma (5.3.2),

$$\| x(\alpha) - \bar{x}(x(\alpha)) \| \leq 2\sigma(Q, A, c, b) \| (-Ax(\alpha) + b)_+ \|.$$

Q.E.D.

It is very interesting that here for the convex quadratic program, we get the same error bound as that for the linear program except for a different constant $\sigma(Q, A, c, b)$. The error bound is simple and easily to compute.

Interior Approach Consider the quadratic program

$$\begin{aligned} &min \quad \tfrac{1}{2} xQx + cx \\ &st \qquad Ax = b \\ &\qquad\quad\; x \geq 0, \end{aligned} \tag{5.3.12}$$

where Q, c, A and b are defined as before and the associated interior penalty function minimization problem

$$\begin{aligned} &min \quad Q(x, \gamma) := \tfrac{1}{2} xQx + cx - \gamma \textstyle\sum_{i=1}^{n} \log x_i \\ &s.t. \qquad\qquad Ax = b \\ &\qquad\qquad\qquad x > 0, \end{aligned} \tag{5.3.13}$$

where $\gamma > 0$ is a constant parameter.

We first establish the following two different error bounds in terms of the KKT violations for (5.3.12) by using different two error residuals $s(x)+s(x)^{\frac{1}{2}}$ and $r(x)+s(x)$. These two error bounds will lead to different algorithm-generated error bounds later on. It is unclear which is the better error bound.

Lemma 5.3.4 *Let (5.3.12) be solvable. Then there exists a $\sigma(Q,A,b,c)$ such that*

$$\begin{aligned}
&\|(x,u)-(\bar{x}(x,u),\bar{u}(x,u))\| \leq \\
&\quad \sigma(Q,A,b,c)\|(-Qx-c+A^Tu,\ -x,\ x(Qx+c-A^Tu)+u(Ax-b))_+, \\
&\quad Ax-b\|_1 + \|(-Qx-c+A^T,\ -x,\ x(Qx+c-A^Tu)+u(Ax-b))_+, \\
&\quad Ax-b\|^{\frac{1}{2}}, \qquad (5.3.14)
\end{aligned}$$

where $(\bar{x}(x,u),\bar{u}(x,u))$ is the orthogonal projection of (x,u) on the set of the KKT pairs of (5.3.12).

Proof. (5.3.12) is equivalent to

$$\begin{aligned}
min \quad & \tfrac{1}{2}xQx+cx \\
st \quad & \bar{A}x \geq \bar{b} \\
& x \geq 0, \qquad (5.3.15)
\end{aligned}$$

where

$$\bar{A}=\begin{bmatrix} A \\ -A \end{bmatrix}, \qquad \bar{b}=\begin{bmatrix} b \\ -b \end{bmatrix}.$$

Since it is a symmetric convex quadratic program, it is equivalent to the LCP of finding some $w \in I\!R^{n+2m}$ such that

$$Mw+q \geq 0, \quad w \geq 0, \quad w(Mw+q)=0 \qquad (5.3.16)$$

where

$$M = \begin{bmatrix} Q & -\bar{A}^T \\ \bar{A} & 0 \end{bmatrix}, \qquad q = \begin{bmatrix} c \\ -\bar{b} \end{bmatrix}.$$

It is easy to see that M is positive semi-definite. Therefore by Theorem 2.3.1, there exists a $\sigma(Q, A, b, c)$ such that $\forall w = (x, u_+, u_-) \in I\!R^{n+2m}$

$$\|w - \bar{w}(w)\| \leq \sigma(Q, A, b, c)(s(w) + s(w)^{\frac{1}{2}}) \tag{5.3.17}$$

where $\bar{w}(w)$ is the orthogonal projection of w on the solution set of the LCP (5.3.16). Now let $(x, u) \in I\!R^{n+m}$ and take $w = (x, u_+, u_-)$ where $u_+ = (u)_+$, $u_- = (-u)_+$. Then for such a w there exists a $\bar{w}(w) = (\bar{x}(w), \bar{u}_+(w), \bar{u}_-(w))$ such that (5.3.17) holds. Let the norm in $s(w)$ indicates 1-norm, then it follows

$$\begin{aligned}
&\|(x, u) - (\bar{x}(x, u), \bar{u}(x, u))\|_1 \\
\leq\ & |(x, u_+ - u_-) - (\bar{x}(w), \bar{u}_+(w) - \bar{u}_-(w))\|_1 \\
\leq\ & \|w - \bar{w}(w)\|_1 \\
\leq\ & \sigma(Q, A, b, c)\|(-Qx + A^T u - c,\ -x,\ x(Qx - A^T u + c) + u(Ax - b))_+, \\
& Ax - b\|_1 + \|(-Qx + A^T u - c,\ -x, x(Qx - A^T u + c) + u(Ax - b))_+, \\
& Ax - b\|_1^{\frac{1}{2}},
\end{aligned}$$

where $(\bar{x}(x, u), \bar{u}(x, u))$ is the projection of (x, u) on the set of the KKT pairs of (5.3.12); the first inequality follows that $(\bar{x}(w), \bar{u}_+(w) - \bar{u}_-(w))$ is a KKT pair; the second follows from the properties of the 1-norm; the third is from (5.3.17). Finally, by the equivalence of norms, (5.3.14) could hold for any norm for some constant $\sigma(Q, A, b, c)$. **Q.E.D.**

Lemma 5.3.5 *Let (5.3.12) be solvable. Then there exists a* $\sigma(Q, A, b, c)$ *such that*

$$\|(x, u) - (\bar{x}(x, u), \bar{u}(x, u))\| \leq \sigma(Q, A, b, c)(s(w) + r(w)) \tag{5.3.18}$$

where $s(w)$ and $r(w)$ are the residuals for LCP defined in (5.3.16) and $(\bar{x}(x,u), \bar{u}(x,u))$ is the orthogonal projection of (x,u) on the set of the KKT pairs of (5.3.12).

Proof. By replacing $s(w)+s(w)^{\frac{1}{2}}$ by $r(w)+s(w)$ in Lemma (5.3.4), the same proof goes through.

By using Lemma 5.3.4, we give the following interior-point-generated error bound on distance between each $x(\gamma)$ and the solution set of the quadratic program in terms of the penalty parameter γ only. The new contribution here is that the bound gives a precise dependence on the penalty parameter γ.

Theorem 5.3.3 *Suppose both (5.3.12) and (5.3.13) are solvable. Let $x(\gamma)$ be a minimizer of (5.3.13) for some $\gamma > 0$, then there exists a $\sigma(Q,A,b,c) > 0$ such that*

$$\|x(\gamma) - \bar{x}(x(\gamma))\| \leq \sigma(Q,A,b,c)(\gamma n + (\gamma n)^{\frac{1}{2}}), \tag{5.3.19}$$

where $(\bar{x}(x(\gamma))$ is the orthogonal projection of $(x(\gamma),u)$ on the set of the KKT pairs of (5.3.12).

Proof. By the KKT optimality conditions, there is a u such that

$$\begin{aligned} Qx(\gamma) + c - \gamma X^{-1}e - A^T u &= 0 \\ Ax(\gamma) &= b \end{aligned}$$

where $X = diag(x(\gamma)), e = (1,1,...,1)$ and $x(\gamma) > 0$. By Lemma 5.3.4, it follows that

$$\begin{aligned} &\|(x(\gamma),u) - (\bar{x}(x(\gamma),u), \bar{u}(x(\gamma),u))\| \\ \leq\ &\sigma(Q,A,b,c)\|(-Qx(\gamma) - c + A^T u,\ -x(\gamma),\ x(\gamma)(Qx(\gamma) + c - A^T u) \end{aligned}$$

$$+u(Ax(\gamma)-b))_+,\ Ax(\gamma)-b\| + \|(-Qx(\gamma)-c+A^Tu,\ -x(\gamma),$$
$$x(\gamma)(Qx(\gamma)+c-A^Tu+u(Ax(\gamma)-b))_+, Ax(\gamma)-b\|^{\frac{1}{2}}$$
$$= \sigma(Q,A,b,c)(\gamma n + (\gamma n)^{1/2})$$

Q.E.D.

Again in (5.3.19) a square root of γ is involved in the bound. Unfortunately, unlike the exterior case, we cannot totally get rid of it by using $r(x)+s(x)$ in Lemma 5.3.5. Instead we obtain another term replacing the square root in the bound which is given in the following theorem. It is unclear which bound is better.

Theorem 5.3.4 *Suppose both (5.3.12) and (5.3.13) are solvable. Let $x(\gamma)$ be a minimizer of (5.3.13) for some $\gamma > 0$, then there exists a $\sigma(Q,A,b,c) > 0$ such that*

$$\|(x(\gamma),u(\gamma) - (\bar{x}(x(\gamma),u(\gamma),\bar{u}(x(\gamma),u(\gamma))\|$$
$$\leq \sigma(Q,A,b,c)(\gamma n + \|min\{x(\gamma),\gamma x(\gamma)^{-1}\}\|), \qquad (5.3.20)$$

where $u(\gamma)$ is an exact multiplier associated with $Ax=b$ in (5.3.13), $x(\gamma)^{-1} = (x_1(\gamma)^{-1}, x_2(\gamma)^{-1},\ldots,x_n(\gamma)^{-1})$ and $(\bar{x}(x(\gamma),u(\gamma)),\bar{u}(x(\gamma),u(\gamma)))$ the orthogonal projection of $(x(\gamma),u(\gamma))$ on the set of the KKT pairs of (5.3.12).

Proof. Similar to the proof of Theorem 5.3.3, we have

$$s(w) = \gamma n.$$

In addition we have

$$r(w) = \|\min\{Mw+q,\ w\}\|$$
$$= \|\min\{\begin{bmatrix} Qx+c-A^Tu \\ Ax-b \\ -Ax+b \end{bmatrix}, \begin{bmatrix} x_+ \\ u_+ \\ u_- \end{bmatrix}\}\|$$

$$= \|\min\{\gamma x(\gamma)^{-1}, x(\gamma)\}\|.$$

Hence by Lemma 5.3.5, the relation (5.3.20) holds.

Augmented Lagrangian Approach Consider the quadratic program (5.3.1) and define the associated augmented Lagrangian as follows:

$$L(x,u,\alpha) := xQx + cx + \frac{1}{2\alpha}[\|(\alpha(b-Ax)+u)_+\|_2^2 - \|u\|_2^2], \quad \alpha > 0. \quad (5.3.21)$$

Let $x(\alpha, u)$ be a minimizer of (5.3.21) with respect to x for fixed $u \geq 0$ and fixed $\alpha > 0$. First, we establish the existence of such an $x(\alpha, u)$.

Lemma 5.3.6 *Let (5.3.1) be solvable. Then (5.3.21) has a minimizer with respect to x for each $\alpha > 0$ and $u \geq 0 \in I\!R^m$.*

Proof. The proof is similar to that of lemma 5.2.3.

The following theorem establishes a bound on the distance between each minimizer $x(\alpha, u)$ of (5.3.21) with respect to x for each fixed $\alpha > 0$ and $u \geq 0$ and the solution set of the quadratic program (5.3.1). Again the corresponding results to the linear program in the last section cannot be directly used here. Instead, the error bounds obtained in the Chapter 2 are used to establish this result. Consequently, the bound we get here are not as sharp as those of the linear case.

Theorem 5.3.5 *Let (5.3.1) be solvable and let $x(\alpha, u)$ be a minimizer of the augmented Lagrangian (5.3.21) respect to x for some $u \geq 0$ and $\alpha > 0$. Assume that all norms are monotonic. Then, there exists $\sigma(A, b, c) > 0$ such that*

$$\begin{aligned}
&\|x(\alpha,u) - \bar{x}(x(\alpha,u))\| \leq \\
&\quad \sigma(Q,A,b,c)(\|(b-Ax(\alpha,u))_+, (\alpha(b-Ax(\alpha,u))+u)_+(Ax(\alpha)-b)_+\| \\
&\quad +\|(b-Ax(\alpha,u))_+, (\alpha(b-Ax(\alpha,u))+u)_+(Ax(\alpha)-b)_+\|^{\frac{1}{2}}) \qquad (5.3.22)
\end{aligned}$$

where $\bar{x}(x(\alpha, u))$ is the orthogonal projection of $x(\alpha, u)$ on the solution set of (5.2.1).

Proof. By Lemma 5.3.6, such a minimizer $x(\alpha, u)$ exists. It satisfies the first order optimality condition

$$\nabla_x L(x(\alpha, u), u, \alpha) = Qx(\alpha, u) + c - A^T[\alpha(b - Ax(\alpha, u)) + u]_+ = 0. \quad (5.3.23)$$

Take u in Lemma 5.3.1 as $[\alpha(b - Ax(\alpha, u)) + u]_+$, then (5.3.14) becomes

$$\begin{aligned} &\|(x, u) - (\bar{x}(x, u), \bar{u}(x, u))\| \leq \\ &\quad \sigma(Q, A, b, c)(\|(b - Ax(\alpha, u), (\alpha(b - Ax(\alpha, u)) + u)_+(Ax(\alpha) - b))_+\| \\ &\quad +\|(b - Ax(\alpha, u), (\alpha(b - Ax(\alpha, u)) + u)_+(Ax - b))_+\|^{\frac{1}{2}}) \\ &\leq \sigma(Q, A, b, c)(\|(b - Ax(\alpha, u))_+, (\alpha(b - Ax(\alpha, u)) + u)_+(Ax(\alpha) - b)_+\| \\ &\quad +\|(b - Ax(\alpha, u))_+, (\alpha(b - Ax(\alpha, u)) + u)_+(Ax(\alpha) - b)_+\|^{\frac{1}{2}}), \end{aligned} \quad (5.3.24)$$

where $x = x(\alpha, u)$. Since

$$\|x(\alpha, u) - \bar{x}(x(\alpha, u))\| \leq \|(x, u) - (\bar{x}(x, u), \bar{u}(x, u))\|.$$

Therefore (5.3.22) holds. **Q.E.D.**

The major difference between the bound here and the bound in (5.2.14) for the linear program is that here there is the additional square toot term. The residual of (5.3.22) shows that any minimizer of the augmented Lagrangian (5.3.21) with respect to x is a solution of the quadratic program (5.3.1) provided it is primal feasible and complementary. However, for small value of the residual the square root dominates and hence the error is not linear in the residual.

The following theorem gives a different error bound for the same distance. The bound does not contain a square root term. However it is not clear if it is better.

Theorem 5.3.6 *Let the quadratic program (5.3.1) be solvable and $x(\alpha,u)$ be a minimizer of the augmented Lagrangian (5.3.21). Assume that all norms are monotonic. Then there exists a $\sigma(Q,A,b,c)>0$ such that*

$$\|x(\alpha,u)-\bar{x}(x(\alpha,u))\| \leq \tag{5.3.25}$$

$$\sigma(Q,A,c,b)(\|(b-Ax(\alpha,u))_+,(\alpha(b-Ax(\alpha,u))+u)_+(Ax(\alpha)-b)_+\|$$

$$+\|\min\{A(\alpha,u)-b,(\alpha(b-Ax(\alpha,u))+u)_+\}\|) \tag{5.3.26}$$

where $\bar{x}(x(\alpha))$ is the orthogonal projection of $x(\alpha)$ on the solution set of (5.3.1).

Proof. In a similar manner to the proof of the previous theorem, we have

$$s(w)=\|(b-Ax(\alpha,u))_+,(\alpha(b-Ax(\alpha,u))+u)_+(Ax(\alpha)-b)_+\|$$

By the definition (5.3.8) of $r(w)$ we further have

$$\begin{aligned} r(w) &= \|\min\{\begin{bmatrix} Qx+c-A^Tu \\ -Qx-c+A^Tu \\ Ax-b \end{bmatrix}, \begin{bmatrix} x_+ \\ x_- \\ u \end{bmatrix}\}\| \\ &= \|\min\{Ax(\alpha,u)-b,(\alpha(b-Ax(\alpha,u)+u)_+\}\|. \end{aligned}$$

Hence, by Lemma (5.3.2),

$$\|x(\alpha,u)-\bar{x}(x(\alpha,u))\| \leq$$

$$\sigma(Q,A,c,b)(\|(b-Ax(\alpha,u))_+,(\alpha(b-Ax(\alpha,u))+u)_+(Ax(\alpha)-b)_+\|$$

$$+\|\min\{A(\alpha,u)-b,(\alpha(b-Ax(\alpha,u))+u)_+\}\|)$$

Q.E.D.

Proximal Point Approach Consider the quadratic program

$$\begin{aligned} min \quad & \tfrac{1}{2}xQx + cx \\ st \quad & Ax \geq b \\ & x \geq 0, \end{aligned} \tag{5.3.27}$$

where Q, c, b and c are defined as before, and the associated proximal point minimization problem

$$\begin{aligned} min \quad & \tfrac{1}{2}xQx + cx + \tfrac{\epsilon}{2}\|x-y\|^2, \text{for fixed } y \\ s.t. \quad & Ax \geq b \\ & x \geq 0, \end{aligned} \tag{5.3.28}$$

where $\epsilon > 0$ is some constant and y is a fixed vector in $I\!R^n$.

Theorem 5.3.7 *Let (5.3.27) be solvable and $x(\epsilon, y)$ be a solution of (5.3.28). Then, there is a $\sigma(Q, A, b, c)$ such that*

$$\begin{aligned} & \|(x(\epsilon), u) - (\bar{x}(x(\epsilon), u), \bar{u}(x(\epsilon), u))\| \\ \leq \; & \sigma(Q, A, b, c)[\|(\epsilon(x(\epsilon) - y), \; -x(\epsilon)(x(\epsilon) - y))_+\| \\ & + \|(\epsilon(x(\epsilon) - y), \; -x(\epsilon)(x(\epsilon) - y))_+\|^{\frac{1}{2}}], \end{aligned} \tag{5.3.29}$$

where u is an exact multiplier associated with $Ax \geq b$ in (5.3.28) and $(\bar{x}(x(\epsilon), u), \bar{u}(x(\epsilon), u))$ is the orthogonal projection of $(x(\epsilon), u)$ on the set of the KKT pairs of (5.3.27).

Proof. Solving (5.3.27) is equivalent to finding $w = (x, u) \in I\!R^{n+m}$ such that it satisfies the KKT optimality condition, that is

$$Mw + q \geq 0, \quad w \geq 0, \quad w(Mw + q) = 0, \tag{5.3.30}$$

where

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

It is easy to see that M is positive semi-definite. By Theorem 2.3.1, there exists a $\sigma(Q, A, b, c)$ such that for $\forall w = (x, u) \in I\!R^{n+m}$

$$\begin{aligned} \|w - \bar{w}(w)\| &\leq \sigma(Q, A, b, c)(s(w) + s(w)^{\frac{1}{2}}) \\ &= \sigma(Q, A, b, c)(\|(-Qx - c + A^T u,\ b - Ax, -x, -u, \\ &\quad x(Qx + c - A^T u) + u(Ax - b))_+\| + \|(-Qx - c - A^T u, \\ &\quad b - Ax, -x, -u,\ x(Qx + c - A^T u) \\ &\quad +u(Ax - b))_+\|^{\frac{1}{2}}, \end{aligned} \tag{5.3.31}$$

where $\bar{w}(w)$ is the orthogonal projection of w on the solution set of (5.3.30). Let $(x, u) \in R^{n+m}$. Now by the assumption that $x(\epsilon)$ solves (5.3.28), there is a multiplier u associated with $x(\epsilon)$ that satisfies the KKT condition

$$Mw + q + v \geq 0, \quad w \geq 0, \quad w(Mw + q + v) = 0,$$

where $w = (x(\epsilon), u)$ and $v = (\epsilon(x(\epsilon) - y), 0) \in R^{n+m}$. Therefore by (5.3.31)

$$\begin{aligned} &\|(x(\epsilon), u) - (\bar{x}(x(\epsilon, u), \bar{u}(x(\epsilon, u))\| \\ &\leq \sigma(Q, A, b, c)\|(\epsilon(x(\epsilon) - y),\ -x(\epsilon)(x(\epsilon) - y))_+\| \\ &\quad +\|(\epsilon(x(\epsilon) - y),\ -x(\epsilon)(x(\epsilon) - y))_+\|^{\frac{1}{2}} \end{aligned}$$

Q.E.D.

5.4 Algorithm-Generated Bounds for Strongly Monotone Nonlinear Complementarity Problems

Exterior Penalty Approach Consider the NCP (3.1.1) and the associated exterior penalty minimization problem

$$\begin{aligned} min \quad & P(x,\alpha) := xF(x) + \frac{\alpha}{2}\|(-F(x))_+\|_2^2, \quad \alpha > 0 \\ s.t. \quad & x \geq 0 \end{aligned} \tag{5.4.1}$$

Theorem 5.4.1 *Let $F(x)$ be differentiable, Lipschitz continuous and strongly monotone on $I\!R^n$. If $x(\alpha)$ solves (5.4.1) for some $\alpha > 0$, then there exists a constant $\sigma > 0$ such that*

$$\|x(\alpha) - \bar{x}\| \leq \sigma\|(-F(x))_+\|, \tag{5.4.2}$$

where $\bar{x}$ is the unique solution of (3.1.1).

Proof. Since $x(\alpha)$ solves (5.4.1), it satisfies the KKT conditions of the problem (5.4.1)

$$\begin{aligned} F(x(\alpha)) + \nabla F(x(\alpha))^T x(\alpha) - \alpha\nabla F(x(\alpha))^T(-F(x(\alpha)))_+ \geq 0, \\ x(\alpha) \geq 0, \\ x(\alpha)(F(x(\alpha)) + \nabla F(x(\alpha))x(\alpha) - \alpha\nabla F(x(\alpha))^T(-F(x(\alpha)))_+) = 0. \end{aligned} \tag{5.4.3}$$

It follows that the point $(x = x(\alpha), u = \alpha(-F(x(\alpha)))_+$ is dual feasible for DP (4.1) of [MaS92] and by property 4.4(i) of [MaS92]

$$\begin{aligned} 0 &\geq -\alpha(-F(x(\alpha)))_+F(x(\alpha)) - x(\alpha)\nabla F(x(\alpha))^T(x(\alpha) - \alpha(-F(x(\alpha)))_+) \\ &= -\alpha(-F(x(\alpha)))_+F(x(\alpha)) + x(\alpha)F(x(\alpha)), \end{aligned}$$

where the second equality is from the complementarity condition of (5.4.3). From the inequality above, we further have

$$(x(\alpha)F(x(\alpha)))_+ \leq (\alpha(-F(x(\alpha)))_+F(x(\alpha)))_+ = 0.$$

Since $(x(\alpha)F(x(\alpha)))_+ = 0$ and $x(\alpha) > 0$, by Lemma 3.2.1,

$$r(x(\alpha)) \leq \|(-F(x(\alpha)))_+\|_1.$$

By Theorem 3.2.1 and equivalence of norms, there exits a constant σ such that

$$\|x(\alpha) - \bar{x}\| \leq \sigma\|(-F(x(\alpha))_+\|.$$

Q.E.D.

Interior Penalty Approach Consider the NCP (3.1.1) and the associated interior penalty function minimization problem

$$\begin{array}{ll} min & Q(x,\gamma) := \frac{1}{2}xF(x) - \gamma\sum_{i=1}^{n}\log F_i(x) \\ s.t. & F(x) > 0 \\ & x \geq 0, \end{array} \tag{5.4.4}$$

where $\gamma > 0$ is a constant parameter.

Lemma 5.4.1 *(Mangasarian and Solodov [MaS92]) Let $F(x)$ be differentiable and monotone on $I\!R^n$, let $\gamma \geq 0$ and $x(\gamma)$ solve (5.4.4). Then*

$$\gamma n \geq x(\gamma)F(x(\gamma)) \geq 0. \tag{5.4.5}$$

Proof. See Proposition 4.6 of [MaS92]. **Q.E.D.**

Theorem 5.4.2 *Let $F(x)$ be differentiable, Lipschitz continuous and strongly monotone on $\mathbb{R}^n$. If $x(\gamma)$ solves (5.4.4). Then there exists a constant σ such that,*

$$|x(\gamma) - \bar{x}\| \leq \sigma(\gamma n)^{\frac{1}{2}}, \tag{5.4.6}$$

where $\bar{x}$ is the unique solution of (3.1.1).

Proof. Since $x(\gamma) \geq 0,\ F(x(\gamma)) > 0$ and by Lemma 5.4.1, $\gamma n \geq x(\gamma)F(x(\gamma)) \geq 0$. By Theorem 3.2.1 and Lemma 3.2.1, there exists a σ such that

$$\|x(\gamma) - \bar{x}\| \leq \sigma(\gamma n)^{\frac{1}{2}}.$$

Q.E.D.

5.5 Error Bounds for Strongly Convex Programs

Let the problem be

$$\begin{aligned} min \quad & f(x) \\ s.t. \quad & g(x) \leq 0 \end{aligned} \tag{5.5.1}$$

where $f(x)$ is differentiable and strongly convex with constant k, that is

$$f(x) - f(y) - \nabla f(y)(x-y) \geq k\|x-y\|^2, \quad \forall x, y \in \mathbb{R}^n,$$

and $g(x)$ is differentiable convex and satisfies some CQ.

We obtain now results for (5.5.1) similar to those of [MaD88], but without any non-negativity on x.

Lemma 5.5.1 *Let $f(x)$ be differentiable and strongly convex with constant k, and $g(x)$ differentiable and convex, let*

$$F(z) := \begin{bmatrix} \nabla f(x) + u\nabla g(x) \\ -g(x) \end{bmatrix} \tag{5.5.2}$$

where $z = (x, u) \in \mathbb{R}^{n+m}$, then for any $z, \bar{z}$ where $u, \bar{u} \geq 0$, we have

$$(F(z) - F(\bar{z}))(z - \bar{z}) \geq k\|x - \bar{x}\|^2 \tag{5.5.3}$$

Proof.

$$\begin{aligned}
&(F(z) - F(\bar{z}))(z - \bar{z}) \\
&= (\nabla f(x) - \nabla f(\bar{x})(x - \bar{x}) + (u\nabla g(x) - \bar{u}\nabla g(\bar{x}))(x - \bar{x}) \\
&\quad -(g(x) - g(\bar{x}))(u - \bar{u}) \\
&\geq k\|x - \bar{x}\|^2 + u[g(\bar{x}) - g(x) - \nabla g(x)(\bar{x} - x)] \\
&\quad +\bar{u}[g(x) - g(\bar{x}) - \nabla g(\bar{x})(x - \bar{x})] \\
&\geq k\|x - \bar{x}\|^2
\end{aligned}$$

Q.E.D.

Exterior Penalty Approach Consider the problem (5.5.1) and define the associated exterior penalty function as follows

$$P(x, \alpha) := f(x) + \frac{\alpha}{2}\|(g(x))_+\|_2^2, \quad \alpha > 0 \tag{5.5.4}$$

Theorem 5.5.1 *Let $x(\alpha)$ be a minimizer of $P(x, \alpha)$ for some $\alpha > 0$. Then we have*

$$\|x(\alpha) - \bar{x}\| \leq k^{-\frac{1}{2}}[-\alpha\|(g(x(\alpha)))_+\|_2^2 + \beta\|(g(x(\alpha)))_+\|]^{\frac{1}{2}}, \tag{5.5.5}$$

where $\bar{x}$ is the unique solution of (5.5.1) and β is some fixed constant.

Proof. First of all, since (5.5.1) is a strongly convex program, such an $x(\alpha)$ always exists. By the KKT conditions for (5.5.1), we have

$$\nabla_x L(\bar{x}, \bar{u}) := \nabla f(\bar{x}) + \bar{u}\nabla g(\bar{x}) = 0,$$
$$\bar{u} \geq 0, \quad g(\bar{x}) \leq 0, \quad \bar{u}g(\bar{x}) = 0.$$

By Lemma 5.5.1, for any $(x, u) \in I\!R^{n+m}$ where $u \geq 0$ and the KKT pair $(\bar{x}, \bar{u})$

$$\begin{aligned}
\|x - \bar{x}\| &\leq k^{-\frac{1}{2}}[(F(z) - F(\bar{z}))(z - \bar{z})]^{\frac{1}{2}} \\
&= k^{-\frac{1}{2}}[(z - \bar{z})F(z) - (z - \bar{z})F(\bar{z})]^{\frac{1}{2}} \\
&= k^{-\frac{1}{2}}[x\nabla_x L(x, u) - ug(x) - \bar{x}\nabla_x L(x, u) + \bar{u}g(x) \\
&\quad --(x - \bar{x})\nabla_x L(\bar{x}, \bar{u}) + g(\bar{x})(u - \bar{u})]^{\frac{1}{2}} \\
&\leq k^{-\frac{1}{2}}[x\nabla_x L(x, u) - ug(x) + \gamma\|\nabla_x L(x, u)\| + \beta\|(g(x))_+\| \\
&\quad -g(\bar{x})\bar{u} + ug(\bar{x})]^{\frac{1}{2}} \\
&\leq k^{-\frac{1}{2}}[x\nabla_x L(x, u) - ug(x) + \tau\|\nabla_x L(x, u)\| \\
&\quad +\beta\|(g(x))_+\|]^{\frac{1}{2}} \qquad (5.5.6)
\end{aligned}$$

where $\tau := \|\bar{x}\|$ and $\beta := inf\{\|u\| \mid u \text{ is any KKT multiplier}\}$. Since $x(\alpha)$ minimizes $P(x, \alpha)$, we have

$$\nabla f(x(\alpha)) + \alpha \sum_{i=1}^{m} (g_i(x(\alpha)))_+ \nabla g_i(x(\alpha)) = 0.$$

Let $u = \alpha(g(x(\alpha)))_+$, then $\nabla_x L(x(\alpha), u) = 0$. Hence (5.5.6) becomes

$$\begin{aligned}
\|x(\alpha) - \bar{x}\| &\leq k^{-\frac{1}{2}}(-ug(x(\alpha)) + \beta\|(g(x(\alpha)))_+\|^{\frac{1}{2}} \\
&\leq k^{-\frac{1}{2}}(-\alpha\|(g(x(\alpha))_+\|_2^2 + \beta\|g(x(\alpha))_+\|)^{\frac{1}{2}}
\end{aligned}$$

Q.E.D.

Interior Penalty Approach Consider the problem (5.5.1) and the associated interior penalty minimization problem for (5.5.1)

$$\begin{aligned} min \quad & Q(x,\gamma) := f(x) - \gamma \textstyle\sum_{i=1}^{m} \log(-g_i(x)) \\ s.t. \quad & g(x) < 0, \end{aligned} \tag{5.5.7}$$

where $\gamma > 0$ is some parameter.

Theorem 5.5.2 *Let (5.5.7) be solvable and $x(\gamma)$ be a solution for some $\gamma > 0$. Then, we have*

$$\|x(\gamma) - \bar{x}\| \leq (\frac{m\gamma}{k})^{\frac{1}{2}} \tag{5.5.8}$$

where $\bar{x}$ is the unique solution of (5.5.1).

Proof. Since $x(\gamma)$ solves (5.5.7), it satisfies the KKT condition

$$\nabla f(x(\gamma)) - \gamma \sum_{i=1}^{m} \frac{\nabla g_i(x(\gamma))}{g_i(x(\gamma))} = 0 \tag{5.5.9}$$

Let $u_i = -\frac{\gamma}{g_i(x(\gamma))}$, then from (5.5.6)

$$\begin{aligned} \|x(\gamma) - \bar{x}\| & \leq (\frac{-ug(x(\gamma))}{k})^{\frac{1}{2}} \\ & \leq (\frac{m\gamma}{k})^{\frac{1}{2}} \end{aligned}$$

Q.E.D.

Chapter 6

Conclusion

6.1 Summary of Work

We have developed new error bounds for linear complementarity problems, quadratic programs, strongly monotone nonlinear complementarity problems and strongly convex programs. For a certain natural residual $r(x)$, it was shown that the residual can serve both as an upper and lower bound to the distance to the solution set of an LCP from the point x. This bound is local for all matrices and global for an R_0 matrix. Moreover, for an approximate solution generated by a computational algorithm such as interior penalty, exterior penalty, proximal point and augmented Lagrangian, we have been able to bound the distance between an approximate solution and the solution set of a given problem. The problems that we have considered include linear programs, convex quadratic programs, strongly monotone nonlinear complementarity problems and strongly convex programs. The bounds involve computable quantities such as constraint violations, penalty parameters and violations of the complementarity condition. These bounds are computable, and hence, serve as a measure of accuracy for the approximate solution computed by the algorithm.

6.2 Some Open Questions Relating to New Error Bounds

Following are some of questions that remain open in the sense that we were unable to either give a proof of the conjecture or construct a counterexample to it. For $t(x)$ defined in (2.4.1), is $t(x) + t(x)^{\frac{1}{2}}$ a global error bound for an LCP(M, q) for a general symmetric M? What is the largest class of LCPs for which $r(x)$ is a error bound? The broader question might be: Are there any local error bounds for convex nonlinear programming problems that are not strongly convex, and for monotone nonlinear complementarity problems that are not strongly monotone? Also, is it possible to obtain error bounds for nonsmooth or nondifferentiable problems? A typical case might be an objective function that is piece-wise differentiable.

6.3 Further Research on Error Bound Applications

So far, the error bound results have been successfully used to analyze the error reduction ratio for various algorithms. As a result, convergence and linear convergence rates have been established for a number of algorithms ([LuT92a] [Ma91] [LuT92b]), the convergence of which was not known previously. Typical examples of these algorithms are the matrix splitting algorithms for solving the nonmonotone symmetric linear complementarity problem ([LuT92a]). It can be further shown that these algorithms with a line search step will also result in the same convergence and linear convergence rate by using an error bound argument. The common denominator to most of the convergence proofs consists of an error bound and a forcing function. The error bound is used to bound the distance between the current iterate and the solution set of the problem by the difference between consecutive iterates. The forcing function forces the difference between the consecutive iterates to approach zero. Using an error bound alone without a proper forcing function may fail to work. This is what happens when we deal with the (even monotone) nonsymmetric linear complementarity problem by using matrix splitting algorithms. The following is an example which shows this situation. Let

$$M = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}, \qquad q = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The solution set $\bar{X}$ of LCP(M, q) is $\{x \mid x_1 = x_2 = x_3 \geq 0\}$. It is a nonsymmetric monotone LCP since $M + M^T = 0$. Consider the matrix splitting with

$$M = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix} = B + C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} -1 & -1 & 1 \\ 1 & -1 & -1 \\ -1 & 1 & -1 \end{bmatrix}.$$

Since

$$B - C = \begin{bmatrix} 2 & 1 & -1 \\ -1 & 2 & 1 \\ 1 & -1 & 2 \end{bmatrix}$$

is positive definite, then $M = B + C$ is a regular splitting. Also B is positive definite. Consider the following regular splitting scheme for solving LCP(M, q).

$$\begin{aligned} x^{i+1} &:= (x^{i+1} - Bx^{i+1} - Cx^i - q)_+ \\ &= (-Cx^i)_+ \end{aligned}$$

Since B is positive definite, the sequence $\{x^i\}$ is well-defined ([CoPS92]). Now take

$$x^0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \text{Then} \quad x^{6k} = 8k \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad k = 1, 2 \ldots.$$

Therefore $\{x^k\}$ diverge. So the regular matrix splitting algorithm fails to solve the nonsymmetric monotone LCP. Here an error bound implies that the distance from the current iterate to the solution set of the LCP is bounded by the difference between consecutive iterates. The problem is that, for nonsymmetric LCP, there is no appropriate quadratic forcing function associated with it. We need some alternative strategies to deal with such cases.

Another way to use an error bound might be as follows. Generate an appropriate objective function to be minimized based on an error bound ([Sh86] [MaS92] [LMRS92]). The advantage is that it is possible to obtain an unconstrained minimization problem for a constrained problem. Also for each iterate, it is also easy to bound the distance between it and the solution set. In addition from the reduction rate of the objective, we can get the reduction rate of the distance between the iterate and the solution set. However generating such a problem by using an error is not easy. In particular, we may lose the differentiability or convexity of the objective function. These are challenging problems for further research.

Bibliography

[BaS79] M. S. Bazaraa and C. M. Shetty: *"Nonlinear Programming Theory and Algorithms"*, Wiley, New York (1979).

[CoPS92] R. W. Cottle, J.-S. Pang & R. E. Stone: *"The Linear Complementarity Problem"*, Academic Press 1992.

[FeM91] M. C. Ferris & O. L. Mangasarian: "Error bounds and strong upper semicontinuity for monotone affine variational inequalities", Technical Report # 1056, Computer Sciences Department, University of Wisconsin, Madison, WI(Nov. 1991).

[Fuk92] M. Fukushima: "Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems", Mathematical Programming 53, 1992, 99-110.

[FraWol56] M. Frank & P. Wolfe: "An algorithm for quadratic programming", Naval Research Logistics Quarterly 3, 95-110.

[FiM68] A. V. Fiacco and G. P. McCormick: *"Nonlinear Programming: Sequential Unconstrained Minimization Techniques"*, Wiley, New York

(1968).

[GHR92] O. Güler, A. J. Hoffman & U. R. Rothblum: "On approximations to solutions of systems of linear inequalities", University of Maryland, Department Mathematics and Statistics Working Paper, (October 1991; revised September 1992).

[Gül91] O. Güler: "Distance to a convex polyhedron and Hoffman's lemma", University of Iowa, Department of Management Sciences Working Paper, (July 1991).

[Hes69] M. R. Hestenes: "Multiplier and gradient methods", J. Optimization Theory and Applications 19, 1969, 141-159.

[Hof52] A. J. Hoffman, "On approximate solutions of systems of linear inequalities", J. Res. Natl. Bur. Standards, 49 1952, 263-265.

[Li91] W. Li: "Lipschitz continuity and strong uniqueness in linear programs and polyhedral approximations", Preprint, Department of Mathematics and Statics, Old Dominion University, Norfolk, VA(1991)

[LMRS92] Z.-Q. Luo, O. L. Mangasarian, J. Ren & M. V. Solodov: "New error bounds for the linear complementarity problem", Technical Report # 1112, Computer Sciences Department, University of Wisconsin, Madison, WI(Sep. 1992), Mathematics of Operations Research, to appear.

[LuT91] Z.-Q. Luo & P. Tseng: "On the convergence rate of dual ascent methods for linearly constrained convex minimization", Mathematics of

Operations Research 18, 1993.

[LuT92a] Z.-Q. Luo & P. Tseng: "Error bound and the convergence analysis of matrix splitting algorithms for the affine variational inequality problem", SIAM J. Optimization 2 (February 1992), 43-54.

[LuT92b] Z.-Q. Luo & P. Tseng: "On the linear convergence of descent methods for convex essentially smooth minimization", SIAM J. Optimization 30, 1992 (March). 408-425.

[LuT92c] Z.-Q. Luo & P. Tseng: "On the convergence of the coordinate descent method for convex differentiable minimization", J. Optimization Theory & Application 72, 1992, 7-35.

[LuT92d] Z.-Q. Luo & P. Tseng: "Error bounds and convergence analysis of feasible descent methods: A general approach", Annuals of Operations Research 32, 1993.

[Ma83] O. L. Mangasarian: "Least-norm linear programming solution as unconstrained minimization problem", J. of Mathematical Analysis and Applications 92, 1983, 240-251.

[Ma86] O. L. Mangasarian: "Some applications of penalty functions in mathematical programming", Lecture Notes in Mathematics No. 1190, "Optimization and Related Fields", R. Conti, E. De Giorgi and F. Giannessi, Editors, Springer-Verlag, Berlin 1986, 307-329.

[Ma90] O. L. Mangasarian: "Error bounds for non-degenerate monotone linear complementarity problems", Mathematical Programming 48, 1990, 437-445.

[Ma91] O. L. Mangasarian: "Global error bounds for monotone affine variational inequality problems", Linear Algebra and Its Applications 174, 1992, 153-163.

[MaD88] O. L. Mangasarian & R. De Leone: "Error bounds for strongly convex programs and (super) linearly convergent iterative schemes for the least 2-norm solution of linear programs", Applied Mathematics and Optimization 17, 1988, 1-14.

[MaM79] O. L. Mangasarian & R. R. Meyer: "Nonlinear perturbation of linear programs", SIAM J. Control and Optimization 17, 1979, 745-752.

[MaPa90] R. Mathias & J.-S. Pang: "Error bounds for the linear complementarity problem with a P-Matrix", Linear Algebra & Its Application 132, 1990, 123-136.

[MaS86] O. L. Mangasarian & T. -H. Shiau: "Error bounds for monotone linear complementarity problems", Mathematical Programming 36, 1986, 81-89.

[MaS87] O. L. Mangasarian & T. -H. Shiau: "Lipschitz continuity of solutions of linear inequalities, programs and complementarity problems", SIAM J. Control & Optimization 25, 1987, 538-595.

[MaS92] O. L. Mangasarian & M. V. Solodov: "Nonlinear complementarity as unconstrained and constrained minimization", Technical Report # 1074, Computer Sciences Department, University of Wisconsin-Madison, WI(Jan. 1992), Mathematical Programming B, to appear.

[Mur88] K. G. Murty: *"Linear Complementarity, Linear and Nonlinear Programming"*, Heldermann Verlag, Berlin (1988).

[Pa86] J.-S. Pang: "Inexact Newton methods for the nonlinear complementarity problem", Mathematical Programming 36,1986, 54-71.

[Pow69] M. J. D. Powell: "A method for nonlinear constraints in minimization problems", Optimization, R. Fletcher, ed., Academic Press, 1969, 283-298.

[Rob73] S. M. Robinson: "Bounds for error in the solution set of a perturbed linear program", Linear Algebra and Its Applications 6, 1973, 69-81.

[Rob81] S. M. Robinson: "Some continuity properties of polyhedral multifunctions", Mathematical Programming Study 14, 1981, 206-214.