DOCTORAL THESIS

First-order affine scaling continuous method for convex quadratic programming

Yue, Hongwei

Date of Award:
2014

First-order Affine Scaling Continuous Method for Convex Quadratic Programming

YUE Hongwei

A thesis submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Principal Supervisor: Prof. LIAO Li-zhi

Hong Kong Baptist University

February 2014

Declaration

I hereby declare that this thesis represents my own work which has been done after registration for the degree of PhD at Hong Kong Baptist University, and has not been previously included in a thesis, dissertation submitted to this or other institution for a degree, diploma or other qualification.

Signature: _______________

Date: February 2014

Abstract

We develop several continuous method models for convex quadratic programming (CQP) problems with different types of constraints. The essence of the continuous method is to construct one ordinary differential equation (ODE) system such that its limiting equilibrium point corresponds to an optimal solution of the underlying optimization problem. All our continuous method models share the main feature of the interior point methods, i.e., starting from any interior point, all the solution trajectories remain in the interior of the feasible regions.

First, we present an affine scaling continuous method model for nonnegativity constrained CQP. Under the boundedness assumption of the optimal set, a thorough study on the properties of the ordinary differential equation is provided, strong convergence of the continuous trajectory of the ODE system is proved. Following the features of this ODE system, a new ODE system for solving box constrained CQP is also presented. Without projection, the whole trajectory will stay inside the box region, and it will converge to an optimal solution. Preliminary simulation results illustrate that our continuous method models are very encouraging in obtaining the optimal solutions of the underlying optimization problems.

For CQP in the standard form, the convergence of the iterative first-order affine scaling algorithm is still open. Under boundedness assumption of the optimal set and nondegeneracy assumption of the constrained region, we discuss the properties of the ODE system induced by the first-order affine scaling direction. The strong convergence of the continuous trajectory of the ODE system is also proved.

Finally, a simple iterative scheme induced from our ODE is presented for finding an optimal solution of nonnegativity constrained CQP. The numerical results illustrate the good performance of our continuous method model with this iterative scheme.

Keywords: ODE; Continuous method; Quadratic programming; Interior point method; Affine scaling.

Acknowledgements

First and foremost, I would like to take this opportunity to express my deepest gratitude to Prof. LIAO Li-zhi, my supervisor, for his constant encouragement and guidance. Without his consistent illuminating instruction and great patience, this thesis could not have reached its present form. His professional quality and good habits in research fields, outstanding character in behaviors, will benefit me for life. I would also like to thank Dr. YANG Zhanwen, for his invaluable advice and helpful comments in my study. I also thank Dr. HU Jianfeng and Dr. SUN Liming for their helpful discussions.

I wish to thank Hong Kong Baptist University for providing quality courses and the whole person education. I am also greatly indebted to the teachers, secretaries and technicians at the Department of Mathematics for their great help.

Finally, I wish to express my gratitude to my wife, LU Huitian, and my parents, for their love and confidence in me all through these years.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

From a general view, there are discrete iterative methods and continuous path methods in the field of computational optimization. Different from the conventional optimization methods, the main feature of the continuous methods is that a continuous path starting from the initial point can be generated. This path eventually will converge to an equilibrium point (or a limit set), which is exactly an optimal solution (or a subset of the optimal solution set) for the underlying optimization problem [42].

In fact, many iterative schemes can be regarded as the discrete realization of certain continuous path. Among the methods for solving the unconstrained problems, the steepest descend method, the Newton method and the power method, all can be taken as typical discretization examples of the corresponding differential system [12]. In solving the constrained linear programming (or quadratic programming), the famous path-following algorithm also uses Newton's method to trace the central path, which is composed of minimizers for the logarithmic barrier family of problems [46, 47, 48, 82].

So far, it has been hard to verify which (discrete or continuous) method takes an advantage in finding the optimal solution in practice, because their performance depends on the type of underlying problems and solvers (hardware devices) to a great extent. Here, we do not attempt to give any answer to this controversial issue either. It is worth the wait that these two types of methods will give full play to their advantages in view of the actual situation. Based on the following considerations, in this thesis we study the continuous methods for convex quadratic programming

(CQP) problems with different types of constraints. For the continuous methods for CQP, there are many new issues involved as listed below.

(a) Many numerical techniques can be used to follow the associated solution flows.

(b) Ordinary differential equation (ODE) may offer better understanding about the convergence conditions for the corresponding discrete method.

(c) The continuous pathes can provide new search directions for discrete methods.

(d) The ODE system can be applied in large scale integrated circuites (LSIC) if the structure of the continuous method is simple. For more results in this field, please see [28, 73, 74, 75, 76, 77, 78, 80].

In the rest of this chapter, we introduce briefly general framework of continuous methods, meanwhile display some important theoretical results for ODE and convex programming.

1.1 Framework of the Continuous Method

The essence of the continuous method is to construct an ODE (dynamical) system, whose limiting equilibrium point is an optimal solution of the underlying optimization problem. All continuous method models can be represented into the following form

$$\begin{cases} E(x(t)) \\ \frac{dx(t)}{dt} = f(x(t)), \ t \geq t_0, \\ x(t_0) = x^0, \end{cases} \tag{1.1}$$

where $t \in R$, $x(t) : R \to R^n$, $E(x) : R^n \to R$ and $f(x) : R^n \to R^n$, t_0 is the initial time, x_0 is the initial point of $x(t)$. $E(x)$ is called the potential (energy) function, whose value is nonincreasing monotonically as $t \to +\infty$, i.e.,

$$\frac{dE}{dt} = (\nabla E(x(t)))^T \frac{dx(t)}{dt} = \nabla E(x(t)))^T f(x(t)) \leq 0, \ \forall t \geq t_0.$$

Thus the convergence (or weak convergence) of the $x(t)$ can be analyzed. Simultaneously, $f(x)$ is required to be locally Lipschitz continuous (or together with other conditions) to ensure that the solution trajectory $x(t)$ is well defined. In other words, there exists a unique solution $x(t)$ for system (1.1) on $[t_0, +\infty)$. For more details about continuous method framework, please see [42]. In some special cases, $f(x)$ may be replaced by $f(x, t)$, thus system (1.1) will be nonautonomous.

To solve constrained optimization problems, there are three ways in structuring the $f(x)$. Here we only list some representative references:

(i) adopting the gradient of certain penalty function [65, 72];

(ii) employing the KKT equations of the underlying optimization problem [71, 85] and;

(iii) taking use of the projection together with variational inequality [74, 75, 79, 83].

Undoubtedly, the three technologies above play important roles in constructing continuous methods (or neural networks) in the past decades. In the case (i), since the equilibrium point of (1.1) corresponds to the minimizer of the penalty function, we only get an approximate solution, which may be very poor. In the case (ii), problem dimension will become larger because the dual variable must be added to the ODE system. In the case (iii), projection technology may bring numerical difficulties, such as the phenomenon of zig-zags.

1.2 Preliminaries

1.2.1 Ordinary Differential Equation

Now, we recall some important definitions and theoretical results in the field of ODE. In our continuous method models, all ODE systems (autonomous dynamical systems) can be stated as

$$\begin{cases} \frac{dx(t)}{dt} = h(x(t)),\ t \geq 0, \\ x(0) = x^0, \end{cases} \tag{1.2}$$

where $x(t) : R \to R^n$, $h(x) : R^n \to R^n$.

Definition 1.1. *If a point $x^* \in R^n$ satisfies $h(x^*) = 0$, then x^* is called an equilibrium point of the ODE system (1.2).*

Definition 1.2. *$h : \mathcal{X} \to \mathcal{Y}$ is said to be Lipschitz continuous if there exists a real constant L such that*

$$\|h(x) - h(y)\| \leq L\|x - y\|,\ \forall\ x, y \in \mathcal{X}.$$

Definition 1.3. *Let $h : \mathcal{D} \to R^n$ be a continuous function defined in the open set $\mathcal{D} \subseteq R^n$. We say that f is locally Lipschitz continuous if for each $x^0 \in \mathcal{D}$, there is an open set $\mathcal{U} \subseteq \mathcal{D}$ containing x^0 such that there is a constant $L > 0$ such that if x, $y \in \mathcal{U}$, then*

$$\|h(x) - h(y)\| \leq L\|x - y\|.$$

The following Lemma is always used to verify whether the solution of ODE (1.2) is well defined on $[0,\ +\infty)$,

Lemma 1.1. *[37] Let $\mathcal{D} \subseteq R^n$ be an open set with $x^0 \in \mathcal{D}$ and $h : \mathcal{D} \to R^n$ be locally Lipschitz continuous with $\mathcal{D}$. Then the ODE system (1.2) has a unique solution $x(t)$ with $[0,\ \alpha)$ being the right maximal interval of existence. Furthermore, if there exists a compact set $\Omega \subseteq \mathcal{D}$ such that the curve*

$$\Gamma_{x^0} = \{x(t) \in R^n \mid t \in [0, \alpha)\} \subseteq \Omega,$$

then $\alpha = +\infty$.

The following Barbalat Lemma will play an important role for the weak convergence of every continuous method throughout this thesis.

Lemma 1.2. *(Barbalat's Lemma) [61] If the differentiable function $g(t)$ has a finite limit as $t \to +\infty$, and $\dot{g}$ is uniformly continuous, then $\dot{g} \to 0$ as $t \to +\infty$.*

1.2.2 Convex Programming

In this thesis, we will focus on CQP, which has many applications that include portfolio selection, constrained least squares, robotics, and sequential quadratic programming approaches to nonlinear programming problems. It is necessary to recall two fundamental theorems in convex programming, due to the important role they paly. The general convex programming is of the following form

$$\min_{x \in \Omega} \; f(x) \tag{1.3}$$

where Ω is a convex subset of R^n, and $f : R^n \to R$ is convex.

Lemma 1.3. *[25] If $g_i : R^n \to R$, $i = 1, \ldots, m$, are concave functions, and if $G = \{x \mid g_i(x) \geq 0, \; i = 1, \ldots, m\}$ is a nonempty bounded set, then for any set of values $\{\epsilon_i\}$, where $\epsilon_i \geq 0$, $i = 1, \ldots, m$, the set*

$$\{x \mid g_i(x) \geq -\epsilon_i, i = 1, \ldots, m\}$$

is bounded.

By Lemma 1.3, if the optimal solution set of problem (1.3) is bounded, then for any point $x^0 \in \Omega$, the level set $\{x \in \Omega \mid f(x) \leq f(x^0)\}$ is bounded.

Lemma 1.4. *[44, 67] Let $f : R^n \to R$ be a convex twice continuously differentiable function. If $f(\cdot)$ is constant on a convex set $\Omega \in R^n$, then $\nabla f(\cdot)$ is constant on Ω.*

Even though the following two lemmas may not be necessarily relevant to convex function, they are also important for this thesis.

Lemma 1.5. *[58] Suppose scalar function $h(t)$ is differentiable on $[t_0, \; T]$ with $h(t_0) = 0$. If there exists an $M > 0$ such that $|\frac{dh}{dt}| \leq M|h(t)|$, $t \in [t_0, \; T]$, then $h(t) = 0$ for all $t \in [t_0, \; T]$.*

Lemma 1.6. *(Inverse Function Theorem) Let $D_2 \subset R^n$ be open and $f : D_2 \to R^n$ be a continuously differentiable function on D_2. If $f'(\alpha)$ is invertible for some $\alpha \in D_2$, then, there exists a neighborhood U of α and a neighborhood V of $\eta := f(\alpha)$ such that f is an invertible function on U.*

Proof. See Theorem 9.24 in [58]. □

The above inverse function theorem will be used to prove the uniqueness of the solution for nonlinear equations in some neighborhood.

1.2.3 Notation

For simplicity, in what follows, unless otherwise specified, R^n_+ denotes the constrained region $\{x \in R^n \mid x \geq 0\}$, $\|\cdot\|$ denotes the 2-norm, x_j denotes the j-th component of a vector x, e denotes the n-dimensional vector of all ones, e_i denotes the unity column vector whose ith component is 1. For each index subset $J \subseteq \{1, \ldots, n\}$, we denote by x_J the vector composed of those components of $x \in R^n$ indexed by $j \in J$. I_n stands for the $n \times n$ identity matrix, $X = diag(x_1, x_2, \ldots, x_n) \in R^{n \times n}$.

1.3 Outline of the Thesis

Through the interior point algorithms for solving CQP, path following and affine scaling algorithms get most of the attentions, the original ideas behind these two algorithms are completely different. The search direction in the path following algorithm is attained by solving the Newton direction associated with the KKT system of equations of some strictly convex programming, whose objective function is a logarithmic barrier function. Thus the iterative points are always close to the central path [39, 47]. But in the affine scaling algorithm, the search direction is obtained by minimizing a first-order linear (or a quadratic) approximation of the objective function over the intersection of the feasible region with the ellipsoid centered at the current point, and the next point is determined by performing a line search along this direction [51, 67, 63]. The convergence of the first-order affine scaling algorithm has not been solved well so far [33].

We will adopt the framework of the continuous method outlined in [42] to construct our ODE systems. The key idea in the continuous method is to formulate an ODE system for each optimization problem such that the limiting equilibrium point of the ODE corresponds to an optimal solution of the underlying optimization problem. In this thesis, the interior point affine scaling directions will be adopted. As a result, the continuous trajectories of ODEs become the central issue in our study.

In Chapter 2, first we present an affine scaling continuous method model for nonnegativity constrained CQP. Under the boundedness assumption of the optimal solution set, a thorough study on the properties of the ordinary differential equation is provided. The strong convergence of the continuous trajectory is proved. Motivated by features of this ODE system, we present a new ODE system for solving box constrained CQP. Without projection, the whole trajectory will stay inside the box region, and it will converge to an optimal solution. Preliminary simulation results illustrate that our continuous method models are encouraging in obtaining the optimal solutions of the optimization problems.

In Chapter 3, for CQP in the standard form, under boundedness assumption of the optimal solution set and nondegeneracy assumption of the constrained region, we discuss the properties of the ODE system induced by the first-order affine scaling direction. The strong convergence of the continuous trajectory is also proved.

Finally, concluding comments and future study are given in Chapter 4. A simple iterative method is presented for solving nonnegativity constrained CQP. Compared with the classical numerical methods for ODE system, our iterative points are not necessarily close to the real solution trajectory. Preliminary experimental results illustrate the good performance of our iterative method.

Chapter 2

First-order Affine Scaling Continuous Method for Nonnegativity and Box Constrained CQP

2.1 Introduction

The general box (bound) constrained CQP is of the following form

$$\begin{aligned} \min \quad & q(x) = \tfrac{1}{2}x^T Q x + c^T x \\ s.t. \quad & l \le x \le u, \end{aligned} \tag{P_1}$$

where c, l and u are given vectors in R^n, $Q = (q_{ij})_{n\times n} \in R^{n\times n}$. We assume throughout this thesis that Q is symmetric and positive semi-definite.

Problem (P_1) frequently arises in numerical analysis applications, optimal control, and subproblems in general nonlinear optimization algorithms. For more details, please see [17, 54] and references therein.

If for any $i \in \{1, \dots, n\}$, l_i is bounded and $u_i = +\infty$, problem (P_1) can be easily converted into the following nonnegativity constrained quadratic programming

$$\begin{aligned} \min \quad & q(x) = \tfrac{1}{2}x^T Q x + c^T x \\ s.t. \quad & x \ge 0. \end{aligned} \tag{P_2}$$

Nonnegativity constrained quadratic programming problems always arise in science, engineering and business, and they may fall into nonnegative least-squares problems. In support vector machines, computing the maximum margin hyperplane also gives rise to a nonnegativity constrained quadratic programming problem [60].

Furthermore, if l_i and u_i are both bounded, and $l_i < u_i$ for any $i \in \{1, \ldots, n\}$, problem (P_1) can be converted into the following special box constrained quadratic programming

$$\begin{aligned} \min \quad & q(x) = \tfrac{1}{2}x^T Qx + c^T x \\ s.t. \quad & 0 \le x \le e. \end{aligned} \tag{P_3}$$

For the interior point algorithms, central path is an important concept. Considering the following logarithmic barrier function optimization problem

$$\begin{aligned} \min \quad & \tfrac{1}{2}x^T Qx + c^T x - \mu \sum_{j=1}^{n} \ln x_j \\ s.t. \quad & x > 0, \end{aligned} \tag{$P_2(\mu)$}$$

where $\mu > 0$ is a barrier penalty parameter. The corresponding KKT system is

$$\begin{cases} Qx + c = s, \;\; s > 0, \\ Xs = \mu e, \;\; x > 0, \end{cases} \tag{2.1}$$

where $s \in R^n$. Under the boundedness assumption of the optimal solution set for problem (P_2), for any given $\mu > 0$, there exists a unique solution of the KKT system (2.1) for x and s in terms of μ, which can be denoted by $x(\mu)$ and $s(\mu)$, respectively. The pair $(x(\mu), s(\mu))$ is called the central point corresponding to μ. Denote central path $\{(x(\mu), s(\mu)) \mid \mu > 0\}$, which has continuous first-order derivatives with respect to μ. After a straightforward calculation, we have

$$\frac{dx}{d\mu} = \frac{1}{\mu^2}(I_n + \frac{1}{\mu}X^2 Q)^{-1} X^2 (Qx + c). \tag{2.2}$$

For analytical simplicity, we define $t = \frac{1}{\mu}$, then (2.2) becomes

$$\frac{dx}{dt} = -(I_n + tX^2 Q)^{-1} X^2 (Qx + c). \tag{2.3}$$

The right-hand-side of (2.3) is the centering direction for problem (P_2). The existence and convergence of the central path for the general convex programming can be found in [52, 23], which provide very detailed discussion.

At some interior point $x > 0$, we consider the following first-order affine scaling subproblem

$$\begin{aligned} \min_{d} \quad & (Qx + c)^T d \\ s.t. \quad & \|X^{-1}d\|^2 \leq \beta^2 < 1, \end{aligned} \tag{2.4}$$

where $\beta > 0$ is a constant. The optimal solution of (2.4) is

$$d = -\frac{\beta X^2(Qx + c)}{\|X(Qx + c)\|}, \tag{2.5}$$

which is a first-order affine scaling direction for problem (P_2).

If $Q = 0$, i.e., the problem (P_2) reduces to linear programming, then the centering direction and the affine scaling direction are the same.

In this chapter, we will adopt the framework of a continuous method outlined in [42]. The key idea in the continuous method is to formulate ODE for each optimization problem such that the limiting equilibrium point of the ODE corresponds to an optimal solution of the underlying optimization problem. In constructing the ODE system for problem (P_2), interior point affine scaling direction will be adopted. As a result, the continuous trajectory becomes the central issue in our study.

The rest of this chapter is organized as follows. In Section 2.2, an ODE will be constructed for problem (P_2). Then, a thorough study on the continuous trajectory of this ODE will be investigated. Various theoretical properties including the strong convergence will be explored. In Section 2.3, similar study and investigations for problem (P_3) will be conducted. Some preliminary numerical results of the two ODEs are illustrated in Section 2.4. Finally, some concluding remarks are drawn in Section 2.5.

2.2 An Affine Scaling Continuous Method for Nonnegativity Constrained CQP

The presentation of this section will be organized in the following three subsections

(1) Construction of an ODE. In addition, some fundamental properties for this ODE will be discussed;

(2) Properties for the continuous trajectory. The focus in this subsection is two folds. First, the weak convergence of this trajectory will be addressed. Second, the optimality property for the limit set of the continuous trajectory will be shown;

(3) Strong convergence. Strong convergence of the continuous trajectory will be proved here.

2.2.1 Construction of an ODE

Following the geometry for linear programming in [6] and the direction d in (2.5), the direction

$$d(x) = -X^2(Qx + c), \tag{2.6}$$

can be adopted. The KKT system of problem (P_2) can be stated as follows

$$\begin{cases} Qx + c = s, \ \ s \geq 0, \\ Xs = 0, \ \ x \geq 0. \end{cases} \tag{2.7}$$

which is a linear complementary problem. According to the analysis above, we adopt the following direction

$$p_1(x) = -X(Qx + c) \tag{2.8}$$

as the interior point affine scaling direction for problem (P_2). Even though direction $d(x)$ in (2.6) and direction $p_1(x)$ in (2.8) are different, yet they are equivalent (by a factor of X). All of our results on $p_1(x)$ also hold for $d(x)$. Based on the interior point affine scaling direction $p_1(x)$ in (2.8), the following ODE for problem (P_2) can be constructed.

$$\begin{cases} \frac{dx(t)}{dt} = -X(Qx + c), \ t \geq 0, \\ x(0) = x^0 > 0. \end{cases} \tag{2.9}$$

It should be noted that if $Q = 0$, our ODE (2.9) is equivalent to the first differential equation on page 515 in [6]. To simplify the following presentation, in the remaining of this thesis, $x(t)$ (or $X(t)$) will be replaced by x (or X) whenever no confusion would occur, throughout this section we make the following assumption.

Assumption 1. *The optimal solution set for problem (P_2) is bounded.*

For given x^0 in ODE (2.9), define level set

$$L_1(x^0) = \{x \in R^n_+ \mid q(x) \leq q(x^0)\}. \tag{2.10}$$

Following Lemma 1.3 in Chapter 1 and Assumption 1, we have

Lemma 2.1. *For problem (P_2), the level set $L_1(x^0)$ is bounded.*

Since $X(Qx + c)$ is continuously differentiable on R^n, obviously $X(Qx + c)$ is locally Lipschitz continuous on R^n. From Lemma 1.1, there exists a unique solution $x(t)$ for ODE (2.9) on the maximal existence interval $[0, \ \beta)$ for some $\beta > 0$.

Theorem 2.1. *Let $x(t)$ be the solution of ODE (2.9) with the maximal existence interval $[0, \ \beta)$. Then $x(t) > 0$ for any $t \in [0, \ \beta)$.*

Proof. We will prove $x(t) > 0$ for any $t \in [0, \ \beta)$ by contradiction. Suppose that there exists a $t^* \in [0, \ \beta)$ and an $i \in \{1, \ldots, n\}$ such that $x_i(t^*) = 0$. Since $x_i(t)$ is continuous on t, let t^* be the minimum t such that $x_i(t) = 0$ for some $i \in \{1, \ldots, n\}$, i.e., $x(t) > 0$ for all $0 \leq t < t^*$.

Let

$$M = \max_{t \in [0, \ t^*]} \|Qx(t) + c\| + 1, \tag{2.11}$$

and

$$t_1 = \max\{t^* - \frac{1}{2M}, 0\},$$

and $\bar{t}$ be the time satisfying

$$x_i(\bar{t}) = \max_{t\in[t_1,\ t^*]} x_i(t) > 0.$$

Notice that

$$\frac{dx(t)}{dt} = -X(Qx + c),$$

we have

$$x_i(t) = x_i(t^*) + \int_t^{t^*} x_i(\tau)(Qx(\tau) + c)_i d\tau.$$

For any $t \in [t_1,\ t^*]$, notice that $x_i(t^*) = 0$ and $x_i(t) \geq 0$, from the above equation, we have

$$\begin{aligned} x_i(t) &\leq M(t^* - t) \max_{\tau\in[t,\ t^*]} x_i(\tau) \\ &\leq M(t^* - t_1) \max_{t\in[t_1,\ t^*]} x_i(t) \\ &= M(t^* - t_1) x_i(\bar{t}) \\ &\leq \frac{1}{2} x_i(\bar{t}). \end{aligned}$$

Since $t \in [t_1, t^*]$ is arbitrary, taking $t = \bar{t}$, then

$$x_i(\bar{t}) \leq \frac{1}{2} x_i(\bar{t}),$$

thus $x_i(\bar{t}) = 0$, which is a contradiction with the definition of $\bar{t}$. □

Suppose $x(t)$ is the solution of ODE (2.9), and $[0,\ \beta)$ is the corresponding maximal existence interval. By Theorem 2.1, $x(t) > 0$ for any $t \in [0,\ \beta)$, thus

$$\frac{dq(x(t))}{dt} = -(Qx + c)^T X(Qx + c) \leq 0,\ \forall t \in [0,\ \beta), \tag{2.12}$$

i.e., $q(x)$ is nonincreasing monotonically along the solution trajectory $x(t)$, so $x(t)$ is contained in the compact level set $L_1(x^0)$. By Lemma 1.1, the following corollary is obvious.

Corollary 2.1. *There exists a unique solution $x(t)$ for ODE (2.9) on $[0,\ +\infty)$, and $x(t) > 0$ for any $t \in [0,\ +\infty)$.*

2.2.2 Properties for the Continuous Trajectory

Now we show the weak convergence of the solution $x(t)$ of ODE (2.9), i.e., the right-hand-side of ODE (2.9) approaching to zero.

Theorem 2.2. *Let $x(t)$ be the solution of ODE (2.9). Then* $\lim_{t\to+\infty} X(Qx+c) = 0$.

Proof. From Corollary 2.1, we know that the unique solution $x(t)$ of ODE (2.9) is always positive on $[0, +\infty)$. Since $\frac{dq(x)}{dt} = -(Qx+c)^T X(Qx+c) \le 0$ and $q(x)$ is bounded below, $q(x)$ has a finite limit along the trajectory $x(t)$. Obviously, $(Qx+c)^T X(Qx+c)$ is continuously differentiable with respect to x, and $x(t)$ is contained in compact level set $L_1(x^0)$. Therefore, there exists a constant $K_1 > 0$ such that

$$\begin{aligned} |\frac{dq(x)}{dt}|_{t=t_1} - \frac{dq(x)}{dt}|_{t=t_2}| &\le K_1\|x(t_1) - x(t_2)\| \\ &= K_1\|\int_{t_1}^{t_2} X(Qx+c)dt\| \\ &\le K_2K_1|t_1 - t_2|, \end{aligned}$$

where $K_2 = \max_{x\in L_1(x^0)} \|X(Qx+c)\|$. Thus $\frac{dq(x)}{dt}$ is uniformly continuous on $[0, +\infty)$. Then Barbalat's Lemma 1.2 ensures that

$$\lim_{t\to+\infty} (Qx+c)^T X(Qx+c) = 0,$$

Since $x(t)$ is bounded and nonnegative, we have

$$\lim_{t\to+\infty} X(Qx+c) = 0.$$

□

Theorem 2.2 ensures the weak convergence of $x(t)$, i.e. $\frac{dx(t)}{dt} \to 0$ as $t \to +\infty$. The results in the following theorems reveal more properties on the trajectory $x(t)$ of ODE (2.9).

Theorem 2.3. *(i) If x is an optimal solution of problem (P_2) and $x > 0$, then $Qx + c = 0$. (ii) Let $x(t)$ be the solution of ODE (2.9). If $X(Qx + c)|_{x=x^0} = 0$, then $x(t) \equiv x^0$ for all $t \geq 0$. Moreover, x^0 is an optimal solution of problem (P_2).*

Proof. If $x > 0$ is an optimal solution of problem (P_2), the KKT system (2.7) implies $X(Qx + c) = 0$. Thus $Qx + c = 0$, (i) is proved.

From ODE (2.9), we know

$$\frac{dx_i(t)}{dt} = -x_i(t)(Qx + c)_i, \ i = 1, \ldots, n.$$

Therefore,

$$|\frac{dx_i(t)}{dt}| \leq |x_i(t)| \|(Qx + c)\| \leq |x_i(t)| \max_{x \in L_1(x^0)} \|Qx + c\|,$$

by Lemma 1.5, we have $x(t) \equiv x^0$ for all $t \geq 0$.

Furthermore, since $x^0 > 0$, and $X(Qx + c)|_{x=x^0} = 0$, we have

$$Qx^0 + c = 0.$$

Clearly, x_0 is an optimal solution of problem (P_2). □

The result in the following theorem indicates that if x^0 is not an optimal solution for problem (P_2), then ODE system (2.9) will never stop in finite time.

Theorem 2.4. *Let $x(t)$ be the solution of ODE (2.9). If $X(Qx + c)|_{t=0} \neq 0$, then $X(Qx + c) \neq 0$ for any $t \geq 0$.*

Proof. Assume that the conclusion is not true. Then there exists a finite time, say $\bar{t} > t_0$, such that $X(Qx+c)|_{t=\bar{t}} = 0$. From the continuity of $X(Qx+c)$, we can assume that $\bar{t}$ is the minimum t such that $X(Qx+c) = 0$. We know that $X(Qx+c)$ is Lipschitz continuous in bounded set $L_1(x^0)$, and let $\bar{L}$ be the corresponding Lipschitz constant and $\delta = \min\{\frac{\bar{t}}{2}, \frac{1}{2\bar{L}}\}$. Then for any t_1, $t_2 \in [\bar{t}-\delta, \bar{t}]$, we have

$$\begin{aligned} & \|X(Qx+c)|_{t=t_1}\| - \|X(Qx+c)|_{t=t_2}\| \\ \leq \; & \|X(Qx+c)|_{t=t_1} - X(Qx+c)|_{t=t_2}\| \\ \leq \; & \bar{L}\|x(t_1) - x(t_2)\| \\ = \; & \bar{L}\| \int_{t_1}^{t_2} X(Qx+c)|_{t=\tau} d\tau \| \\ \leq \; & \bar{L} \cdot \delta \cdot \max_{\tau \in [\bar{t}-\delta, \bar{t}]} \|X(Qx+c)|_{t=\tau}\|. \end{aligned}$$

Notice that the above inequality is true for any t_1, $t_2 \in [\bar{t}-\delta, \bar{t}]$ and $X(Qx+c)|_{t=\bar{t}} = 0$, then we can choose t_1 and t_2 such that

$$\|X(Qx+c)|_{t=t_1}\| = \max_{\tau \in [\bar{t}-\delta, \bar{t}]} \|X(Qx+c)|_{t=\tau}\|,$$

and

$$\|X(Qx+c)|_{t=t_2}\| = \min_{\tau \in [\bar{t}-\delta, \bar{t}]} \|X(Qx+c)|_{t=\tau}\|,$$

thus we have

$$0 = \min_{\tau \in [\bar{t}-\delta, \bar{t}]} \|X(Qx+c)|_{t=\tau}\| \geq (1-\bar{L}\delta) \max_{\tau \in [\bar{t}-\delta, \bar{t}]} \|X(Qx+c)|_{t=\tau}\|,$$

this implies that $X(Qx+c)|_{t=\tau} = 0$ for any $\tau \in [\bar{t}-\delta, \bar{t}]$ which contradicts with the definition of $\bar{t}$. Thus the proof is complete. □

In the remaining part of this section, we will show that any cluster point of $x(t)$, which is the solution of ODE (2.9), is an optimal solution for problem (P_2). But first, let us define the limit set

$$\Omega_1(x^0) = \{y \in R^n_+ \mid \; y \text{ is a cluster point of } x(t) \text{ of ODE (2.9)}\}. \tag{2.13}$$

Because of the boundedness of $x(t)$, $\Omega_1(x^0)$ is nonempty, compact, and connected (see Theorem 1.1 on page 390 in [15]).

Following the KKT conditions in (2.7) for problem (P_2), we can define the dual estimate as

$$s(x) = Qx + c. \tag{2.14}$$

Note: $s(x(t))$ may be not nonnegative for all $t \geq 0$. Furthermore, we choose an $\bar{x} \in \Omega_1(x^0)$, and define

$$\bar{s} = Q\bar{x} + c. \tag{2.15}$$

Corollary 2.2. *(i) $q(x) = q(\bar{x})$ $\forall x \in \Omega_1(x^0)$. (ii) $Xs(x) = 0$ $\forall x \in \Omega_1(x^0)$, where $s(x)$ is defined in (2.14).*

Proof. (i) Since $\frac{dq(x)}{dt} = -(Qx+c)^T X(Qx+c) \leq 0$ and $q(x)$ is bounded below, it is easy to see that $q(x)$ equals a constant for any $x \in \Omega_1(x^0)$.

(ii) From Theorem 2.2, this is straightforward. □

For the pair $\bar{x}$ and $\bar{s}$, we define

$$\bar{J} = \{j | \bar{s}_j = 0,\ j \in \{1, \cdots, n\}\}, \quad \bar{J}^c = \{1, \ldots, n\} \setminus \bar{J}, \tag{2.16}$$

$$\bar{\Lambda}_1 = \{x \in R^n_+ | x_{\bar{J}^c} = 0,\ q(x) = q(\bar{x})\}. \tag{2.17}$$

From Theorem 2.2, we have

$$\bar{X}\bar{s} = 0 \quad \text{or} \quad \bar{x}_i \bar{s}_i = 0,\ i = 1, \cdots, n.$$

This and the definition of $\bar{J}^c$ imply for any $j \in \bar{J}^c$

$$\bar{s}_j \neq 0 \quad \text{and} \quad \bar{x}_j = 0.$$

This and Corollary 2.2 (i) ensure that the set $\bar{\Lambda}_1$ is nonempty since $\bar{x} \in \bar{\Lambda}_1$. In addition, it is easy to see that $\bar{\Lambda}_1$ is closed. Next we will reveal some properties for $\bar{\Lambda}_1$.

Lemma 2.2. *$\bar{\Lambda}_1$ is convex.*

Proof. Let x be an arbitrary point in the convex hull $co(\bar{\Lambda}_1)$ of $\bar{\Lambda}_1$, i.e., x is a positive linear convex combination of some points in $\bar{\Lambda}_1$. From the definition of $\bar{\Lambda}_1$ in (2.17), we know that $x_{\bar{J}^c}$=0 and $x \geq 0$. From the convexity of $q(x)$, the following inequality holds

$$q(x) \leq q(\bar{x}).$$

On the other hand, let $\Delta x = x - \bar{x}$. Then $(\Delta x)_{\bar{J}^c} = 0$, $\bar{s}^T(\Delta x) = 0$, again by the convexity of $q(x)$, we have

$$\begin{aligned} q(x) &\geq q(\bar{x}) + \nabla q(\bar{x})^T(\Delta x) \\ &= q(\bar{x}) + \bar{s}^T(\Delta x) \\ &= q(\bar{x}). \end{aligned}$$

So $q(x) = q(\bar{x})$ for all $x \in co(\bar{\Lambda}_1)$, thus $x \in \bar{\Lambda}_1$ and $\bar{\Lambda}_1$ is convex. □

Lemma 2.3. $s(x) = \bar{s}$ *for all* $x \in \bar{\Lambda}_1$.

Proof. From the definition of $\bar{\Lambda}_1$, $q(x) = q(\bar{x})$ $\forall x \in \bar{\Lambda}_1$. Then Lemma 2.2 and Lemma 1.4 ensures the result. □

Theorem 2.5. $\Omega_1(x^0) \subseteq \bar{\Lambda}_1$.

Proof. Our proof here is similar to the one for Lemma 8 in [67]. If $\bar{J}^c$ is empty, $\bar{\Lambda}_1$ becomes $\{x \in R^n_+ | q(x) = q(\bar{x})\}$. From Corollary 2.2 (i), the result holds clearly. Suppose there exists a point $\hat{x} \in \Omega_1(x^0)$ but $\hat{x} \notin \bar{\Lambda}_1$ with $\hat{x}_{\hat{j}} > 0$ for some $\hat{j} \in \bar{J}^c$, then $q(\hat{x}) = q(\bar{x})$ and $\hat{x} \geq 0$. Clearly $\bar{\Lambda}_1$ lies inside the bounded level set $L_1(x^0)$, this and $\bar{\Lambda}_1$ being closed ensure that $\bar{\Lambda}_1$ is compact. Thus $s(x)$ is uniformly continuous over $\bar{\Lambda}_1$. Lemma 2.3 implies that, for all $\delta > 0$ sufficiently small, we have

$$|s_j(x)| \geq |\bar{s}_j|/2 \;\; \forall j \in \bar{J}^c, \; \forall x \in U(\bar{\Lambda}_1, \delta), \tag{2.18}$$

where $U(\bar{\Lambda}_1, \delta)$ is the δ-neighborhood of set $\bar{\Lambda}_1$. We take δ small enough so that $\delta < \hat{x}_{\hat{j}}$. Then $\hat{x} \notin U(\bar{\Lambda}_1, \delta)$ since $|\hat{x}_{\hat{j}} - x_{\hat{j}}| = \hat{x}_{\hat{j}} > \delta$ for all $x \in \bar{\Lambda}_1$. Notice $\bar{x} \in \Omega_1(x^0) \cap \bar{\Lambda}_1$

and $\hat{x} \in \Omega_1(x^0)$ but $\hat{x} \notin U(\bar{\Lambda}_1, \delta)$, by the connectivity of $\Omega_1(x^0)$, there must exist a $\tilde{x} \in \Omega_1(x^0) \cap U(\bar{\Lambda}_1, \delta)$ but $\tilde{x} \notin \bar{\Lambda}_1$. $\tilde{x} \in \Omega_1(x^0)$ ensures

$$\tilde{x} \geq 0, \; q(\tilde{x}) = q(\bar{x}).$$

$\tilde{x} \notin \bar{\Lambda}_1$ indicates that there must exist some $r \in \bar{J}^c$ such that $\tilde{x}_r \neq 0$. (2.18) and $\tilde{x} \in U(\bar{\Lambda}_1, \delta)$ imply $|s_j(\tilde{x})| \geq |\bar{s}_j|/2$ for all $j \in \bar{J}^c$, thus $\tilde{X}s(\tilde{x}) \neq 0$, which contradicts with the fact $\tilde{X}s(\tilde{x}) = 0$ since $\tilde{x} \in \Omega_1(x^0)$ from Corollary 2.2 (ii). □

Theorem 2.6. *If $x(t)$ is the solution of ODE (2.9), $\lim_{t \to +\infty} (Qx(t) + c) = \bar{s}$ and $\bar{s} \geq 0$.*

Proof. Based on the continuity of $s(x(t))$, compactness of $\Omega_1(x^0)$, Lemma 2.3 and Theorem 2.5, it is easy to have

$$\lim_{t \to +\infty} (Qx(t) + c) = \bar{s}.$$

Suppose there exists some $\bar{j} \in \{1, \ldots, n\}$ such that $\bar{s}_{\bar{j}} < 0$. For any cluster point $\hat{x} \in \Omega_1(x^0)$, from Corollary 2.2 (ii), we have $\hat{X}s(\hat{x}) = 0$. This, Lemma 2.3, and Theorem 2.5 imply $\hat{X}\bar{s} = 0$, thus $\hat{x}_{\bar{j}} = 0$. Since $s(x(t))$ is continuous, there exists some t_K such that $s_{\bar{j}}(x(t)) < 0$ for all $t \geq t_K$, notice that

$$\frac{dx(t)}{dt} = -X(Qx + c),$$

and $x(t) > 0$ for all $t \geq 0$. We have $\frac{dx_{\bar{j}}(t)}{dt} \geq 0$ and $x_{\bar{j}}(t) \geq x_{\bar{j}}(t_K) > 0$ for all $t \geq t_K$, which contradicts with $\hat{x}_{\bar{j}} = 0$, thus the proof is complete. □

Theorem 2.7. *Any point $x \in \Omega_1(x^0)$ is an optimal solution of problem (P_2).*

Proof. For any $x \in \Omega_1(x^0)$, by Corollary 2.2 (ii), Lemma 2.3, Theorem 2.5, and Theorem 2.6, the following conditions hold

$$\begin{cases} Qx + c = \bar{s}, \; \bar{s} \geq 0, \\ X\bar{s} = 0, \; x \geq 0, \end{cases} \tag{2.19}$$

which are exactly the KKT conditions (2.7). □

The result in Theorem 2.7 ensures that any limit point of $x(t)$ of ODE (2.9) is an optimal solution for problem (P_2).

Theorem 2.8. *Let $x(t)$ be the solution of ODE (2.9). If $\bar{J}^c$ defined in (2.16) is nonempty, there exist two positive constants γ and $\bar{M}$ such that $x_j(t) \leq \bar{M}e^{-\gamma t}$, for all $j \in \bar{J}^c$.*

Proof. If $x(t)$ is the solution of ODE (2.9), we know from Theorem 2.6 that $\lim_{t\to+\infty}[Qx(t)+c] = \bar{s}$ and $\bar{s} \geq 0$. If $\bar{J}^c$ is nonempty, let

$$\gamma = \frac{1}{2}\min_{j\in\bar{J}^c}\{\bar{s}_j\}.$$

Clearly $\gamma > 0$, and there exists a T sufficiently large such that

$$[Qx(t) + c]_j \geq \gamma, \ \forall j \in \bar{J}^c \text{ and } t \geq T.$$

Thus $\forall j \in \bar{J}^c$ and $t \geq T$,

$$\begin{aligned}\frac{d(x_j(t)e^{\gamma t})}{dt} &= e^{\gamma t}(\frac{dx_j(t)}{dt} + \gamma x_j(t)) \\ &= e^{\gamma t}x_j(t)(-[Qx(t) + c]_j + \gamma) \\ &\leq 0.\end{aligned}$$

The above inequality implies that $x_j(t)e^{\gamma t}$ ($\forall j \in \bar{J}^c$) is nonincreasing monotonically when $t > T$, there exists an $\bar{M} > 0$ such that

$$x_j(t)e^{\gamma t} \leq \bar{M}, \ \forall j \in \bar{J}^c,$$

or

$$x_j(t) \leq \bar{M}e^{-\gamma t}, \ \forall j \in \bar{J}^c.$$

□

Theorem 2.8 indicates that if $\bar{J}^c$ is nonempty, $x_j(t)$ ($j \in \bar{J}^c$) will converge to zero exponentially.

2.2.3 Strong Convergence

The task of this section is to prove that the solution $x(t)$ of ODE (2.9) converges to an optimal solution of problem (P_2) as $t \to +\infty$. Our strategy is to show that the set $\Omega_1(x^0)$ defined in (2.13) contains a single point.

Theorem 2.9. $\Omega_1(x^0)$ *defined in (2.13) only contains a single point.*

Proof. Since $\Omega_1(x^0)$ is nonempty, assume $x^* \in \Omega_1(x^0)$ and the number of its nonzero components is maximum for all $x \in \Omega_1(x^0)$, then the index set $\{1, \ldots, n\}$ can be partitioned into two disjoint sets B and N such that

$$B = \{i | x_i^* > 0,\ i \in \{1, \ldots, n\}\} \quad \text{and} \quad N = \{i | x_i^* = 0,\ i \in \{1, \ldots, n\}\}.$$

If $B = \emptyset$, we can conclude there is a single point $x^* = 0$ in $\Omega_1(x^0)$. So we will focus on the case that B is nonempty, without loss of generality, we assume

$$B = \{1, \ldots, k\}\ (k \geq 1) \text{ and } N = \{k+1, \ldots, n\}.$$

Correspondingly, for any $x \in R^n$, it can be denoted by $x = \begin{pmatrix} x_B \\ x_N \end{pmatrix}$. Similarly, we can partition $s = \begin{pmatrix} s_B \\ s_N \end{pmatrix}$, $c = \begin{pmatrix} c_B \\ c_N \end{pmatrix}$ respectively, where x_B, s_B, $c_B \in R^k$, and x_N, s_N, $c_N \in R^{(n-k)}$.

Let $\delta_1 = \frac{1}{2} \min_{i \in B} \{x_i^*\}$. Together with the definition of x^*, we know

$$x_B > 0 \text{ and } x_N = 0\ \forall x \in \Omega_1(x^0) \cap U(x^*, \delta_1), \tag{2.20}$$

where $U(x^*, \delta_1)$ is the δ_1-neighborhood of x^*. Next we will prove that x^* is an isolated point of $\Omega_1(x^0)$.

For any point $x \in \Omega_1(x^0)$, from Lemma 2.3 and Theorem 2.5, we have

$$Qx + c = Qx^* + c \doteq s^*. \tag{2.21}$$

For the convenience of discussion, we denote

$$Q = [q_1, q_2, \ldots, q_n] \text{ and } b = s^* - c.$$

If $x \in \Omega_1(x^0) \cap U(x^*, \delta_1)$, we have $x_B > 0$ and $x_N = 0$. Then the first equality in (2.21) can be written as

$$x_1 q_1 + \cdots + x_k q_k = b \quad \forall x \in \Omega_1(x^0) \cap U(x^*, \delta_1). \tag{2.22}$$

Thus, $rank[q_1, q_2, \ldots, q_k] = rank[q_1, q_2, \ldots, q_k, b]$.

If $rank[q_1, q_2, \ldots, q_k] = k$, b can be expressed uniquely as a linear combination of $q_1, q_2, \ldots, q_k$, thus except for x^*, there is no x in $\Omega_1(x^0) \cap U(x^*, \delta_1)$ such that (2.22) holds, in other words, x^* is an isolated point of $\Omega_1(x^0)$.

If $rank[q_1, q_2, \ldots, q_k] = r < k$ and $r = 0$, then $q_1 = q_2 = \cdots = q_k = 0$, $\frac{dx_i(t)}{dt}$ in ODE (2.9) will be reduced to

$$\frac{dx_i(t)}{dt} = -c_i x_i, \quad i = 1, \ldots, k.$$

From Theorem 2.2, $\lim_{t \to +\infty} \frac{dx_i(t)}{dt} = 0$, thus there are two cases

(a) $c_i = 0 \Rightarrow x_i(t) \equiv x_i^0$, since $x_i^0 > 0$ is arbitrary, therefore the optimal solution set is unbounded, which contradicts with the boundedness of the optimal solution set;

(b) $c_i \neq 0 \Rightarrow x_i^* = 0$, this is a contradiction with the assumption that x_i^* is positive.

So we only consider the case that $1 \leq r < k$, and assume $\{q_{p_1}, q_{p_2}, \ldots, q_{p_r}\}$ is a maximum linearly independent subset of $\{q_1, q_2, \ldots, q_k\}$, and $\{q_{p_{r+1}}, q_{p_{r+2}}, \ldots, q_{p_k}\} = \{q_1, q_2, \ldots, q_k\} \setminus \{q_{p_1}, q_{p_2}, \ldots, q_{p_r}\}$. Thus there exists a matrix $W = (w_{ij})_{(k-r)\times r} \in R^{(k-r)\times r}$ such that

$$q_{p_{r+i}} = \sum_{j=1}^{r} w_{ij} q_{p_j}, \; i = 1, \ldots, k-r. \tag{2.23}$$

We consider the following sub-system (k rows) of the first equation in (2.21)

$$\begin{cases} q_{p_1}^T x = b_{p_1}, \\ \quad\vdots \\ q_{p_r}^T x = b_{p_r}, \\ q_{p_{r+1}}^T x = b_{p_{r+1}}, \\ \quad\vdots \\ q_{p_k}^T x = b_{p_k}. \end{cases} \tag{2.24}$$

Combining (2.24) with (2.23), we have

$$b_{p_{r+i}} = \sum_{j=1}^{r} w_{ij} b_{p_j}, \ i = 1, \ldots, k-r. \tag{2.25}$$

From Corollary 2.2 (ii), we have $X^* s^* = 0$ which implies $s_B^* = 0$. Thus

$$c_B = -b_B,$$

where $b = \begin{pmatrix} b_B \\ b_N \end{pmatrix}$. This and (2.25) indicate

$$c_{p_{r+i}} = \sum_{j=1}^{r} w_{ij} c_{p_j}, \ i = 1, \ldots, k-r. \tag{2.26}$$

Clearly x^* is a solution of (2.24), but linear system (2.24) is degenerate. In overcoming the difficulty caused by the degeneracy of linear system (2.24), we define

$$y_i(t) = \sum_{j=1}^{r} w_{ij} \ln x_{p_j}(t) - \ln x_{p_{r+i}}(t), \quad i = 1, \ldots, k-r, \ t \geq 0, \tag{2.27}$$

where $x(t)$ is the solution of ODE (2.9). From Theorem 2.1, we know $x(t) > 0$ for all $t \geq 0$. Therefore, $y_i(t), \ i = 1, \ldots, k-r$ are well defined for $t \geq 0$. Notice that

$$\frac{dx(t)}{dt} = -X(Qx + c),$$

then we have

$$\begin{aligned} \frac{dy_i(t)}{dt} &= \sum_{j=1}^{r} w_{ij} \frac{\frac{dx_{p_j}(t)}{dt}}{x_{p_j}} - \frac{\frac{dx_{p_{r+i}}(t)}{dt}}{x_{p_{r+i}}} \\ &= \Big(-\sum_{j=1}^{r} w_{ij} q_{p_j}^T + q_{p_{r+i}}^T\Big)x - \sum_{j=1}^{r} w_{ij} c_{p_j} + c_{p_{r+i}} \quad ((2.23) \text{ and } (2.26)) \\ &\equiv 0, \quad i = 1, \ldots, k-r, \ t \geq 0. \end{aligned}$$

Thus there exist $k-r$ constants $\bar{c}_i$ $(i=1,\dots,k-r)$ such that

$$y_i(t) \equiv \bar{c}_i, \quad i=1,\dots,k-r,\ t \geq 0.$$

In particular,

$$\sum_{j=1}^{r} w_{ij} \ln x_{p_j} - \ln x_{p_{r+i}} \equiv \bar{c}_i, \quad i=1,\dots,k-r,\ \forall x \in \Omega_1(x^0) \cap U(x^*,\delta_1). \tag{2.28}$$

Let $H \in R^{r \times k}$ be a matrix generated by choosing r linearly independent rows, say $l_1,\dots,l_r$, from matrix $[q_{p_1},\dots,q_{p_r},q_{p_{r+1}},\dots,q_{p_k}]$. Thus H can be written as

$$H = \begin{pmatrix} q_{l_1p_1} & \cdots & q_{l_1p_r} & q_{l_1p_{r+1}} & \cdots & q_{l_1p_k} \\ q_{l_2p_1} & \cdots & q_{l_2p_r} & q_{l_2p_{r+1}} & \cdots & q_{l_2p_k} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ q_{l_rp_1} & \cdots & q_{l_rp_r} & q_{l_rp_{r+1}} & \cdots & q_{l_rp_k} \end{pmatrix}.$$

Then the following nonlinear system is introduced

$$\begin{cases} q_{l_1p_1}z_1 + \cdots + q_{l_1p_r}z_r + q_{l_1p_{r+1}}z_{r+1} + q_{l_1p_{r+2}}z_{r+2} + \cdots + q_{l_1p_k}z_k = b_{l1}, \\ q_{l_2p_1}z_1 + \cdots + q_{l_2p_r}z_r + q_{l_2p_{r+1}}z_{r+1} + q_{l_2p_{r+2}}z_{r+2} + \cdots + q_{l_2p_k}z_k = b_{l2}, \\ \quad\vdots \\ q_{l_rp_1}z_1 + \cdots + q_{l_rp_r}z_r + q_{l_rp_{r+1}}z_{r+1} + q_{l_rp_{r+2}}z_{r+2} + \cdots + q_{l_rp_k}z_k = b_{lr}, \\ w_{11}\ln z_1 + \cdots + w_{1r}\ln z_r - \ln z_{r+1} - 0 - 0 - \cdots - 0 = \bar{c}_1, \\ w_{21}\ln z_1 + \cdots + w_{2r}\ln z_r - 0 - \ln z_{r+2} - 0 - \cdots - 0 = \bar{c}_2, \\ \quad\vdots \\ w_{(k-r)1}\ln z_1 + \cdots + w_{(k-r)r}\ln z_r - 0 - 0 - 0 - \cdots - \ln z_k = \bar{c}_{k-r}. \end{cases} \tag{2.29}$$

From (2.24) and (2.28), we know that for any $x \in \Omega_1(x^0) \cap U(x^*,\delta_1)$, $z = (x_{p_1},\dots,x_{p_r},x_{p_{r+1}},\dots,x_{p_k})^T$ is a solution of system (2.29).

The Jacobian matrix of nonlinear system (2.29) is

$$J(z) = \begin{pmatrix} q_{l_1p_1} & \cdots & q_{l_1p_r} & q_{l_1p_{r+1}} & q_{l_1p_{r+2}} & \cdots & q_{l_1p_k} \\ q_{l_2p_1} & \cdots & q_{l_2p_r} & q_{l_2p_{r+1}} & q_{l_2p_{r+2}} & \cdots & q_{l_2p_k} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ q_{l_rp_1} & \cdots & q_{l_rp_r} & q_{l_rp_{r+1}} & q_{l_rp_{r+2}} & \cdots & q_{l_rp_k} \\ w_{11}\frac{1}{z_1} & \cdots & w_{1r}\frac{1}{z_r} & -\frac{1}{z_{r+1}} & 0 & \cdots & 0 \\ w_{21}\frac{1}{z_1} & \cdots & w_{2r}\frac{1}{z_r} & 0 & -\frac{1}{z_{r+2}} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{(k-r)1}\frac{1}{z_1} & \cdots & w_{(k-r)r}\frac{1}{z_r} & 0 & 0 & \cdots & -\frac{1}{z_k} \end{pmatrix}.$$

From (2.23), after a series of Gaussian eliminations, the Jacobian matrix $J(z)$ can be converted into

$$\begin{aligned} \bar{J}(z) &= \begin{pmatrix} q_{l_1p_1} & \cdots & q_{l_1p_r} & 0 & 0 & \cdots & 0 \\ q_{l_2p_1} & \cdots & q_{l_2p_r} & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ q_{l_rp_1} & \cdots & q_{l_rp_r} & 0 & 0 & \cdots & 0 \\ w_{11}\frac{1}{z_1} & \cdots & w_{1r}\frac{1}{z_r} & -\frac{1}{z_{r+1}} - \sum_{j=1}^{r} \frac{w_{1j}^2}{z_j} & -\sum_{j=1}^{r} \frac{w_{2j}w_{1j}}{z_j} & \cdots & -\sum_{j=1}^{r} \frac{w_{(k-r)j}w_{1j}}{z_j} \\ w_{21}\frac{1}{z_1} & \cdots & w_{2r}\frac{1}{z_r} & -\sum_{j=1}^{r} \frac{w_{1j}w_{2j}}{z_j} & -\frac{1}{z_{r+2}} - \sum_{j=1}^{r} \frac{w_{2j}^2}{z_j} & \cdots & -\sum_{j=1}^{r} \frac{w_{(k-r)j}w_{2j}}{z_j} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{(k-r)1}\frac{1}{z_1} & \cdots & w_{(k-r)r}\frac{1}{z_r} & -\sum_{j=1}^{r} \frac{w_{1j}w_{(k-r)j}}{z_j} & -\sum_{j=1}^{r} \frac{w_{2j}w_{(k-r)j}}{z_j} & \cdots & -\frac{1}{z_k} - \sum_{j=1}^{r} \frac{w_{(k-r)j}^2}{z_j} \end{pmatrix} \\ &\doteq \begin{pmatrix} P_1 & 0 \\ P_2 & P_3 \end{pmatrix}, \end{aligned}$$

where $P_1 \in R^{r\times r}$, $P_2 \in R^{(k-r)\times r}$, and $P_3 \in R^{(k-r)\times(k-r)}$. Notice the following two facts: (i) $rank(P_1) = r$; and (ii)

$$P_3 = -\begin{pmatrix} \frac{w_{11}}{\sqrt{z_1}} & \frac{w_{12}}{\sqrt{z_2}} & \cdots & \frac{w_{1r}}{\sqrt{z_r}} \\ \frac{w_{21}}{\sqrt{z_1}} & \frac{w_{22}}{\sqrt{z_2}} & \cdots & \frac{w_{2r}}{\sqrt{z_r}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{w_{(k-r)1}}{\sqrt{z_1}} & \frac{w_{(k-r)2}}{\sqrt{z_2}} & \cdots & \frac{w_{(k-r)r}}{\sqrt{z_r}} \end{pmatrix} \begin{pmatrix} \frac{w_{11}}{\sqrt{z_1}} & \frac{w_{12}}{\sqrt{z_2}} & \cdots & \frac{w_{1r}}{\sqrt{z_r}} \\ \frac{w_{21}}{\sqrt{z_1}} & \frac{w_{22}}{\sqrt{z_2}} & \cdots & \frac{w_{2r}}{\sqrt{z_r}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{w_{(k-r)1}}{\sqrt{z_1}} & \frac{w_{(k-r)2}}{\sqrt{z_2}} & \cdots & \frac{w_{(k-r)r}}{\sqrt{z_r}} \end{pmatrix}^T - \begin{pmatrix} \frac{1}{z_{r+1}} & 0 & \cdots & 0 \\ 0 & \frac{1}{z_{r+2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{z_k} \end{pmatrix}.$$

Therefore, $\bar{J}(z)$ and $J(z)$ are invertible if $z\ (\in R^k)\ > 0$.

Now let $F(z) = 0$ be system (2.29). From previous discussions, we know (i) $\forall x \in \Omega_1(x^0) \cap U(x^*, \delta_1)$, $z = (x_{p_1}, \ldots, x_{p_r}, x_{p_{r+1}}, \ldots, x_{p_k})^T$ is a solution of $F(z) = 0$, in particular, $z^* = (x^*_{p_1}, \ldots, x^*_{p_r}, x^*_{p_{r+1}}, \ldots, x^*_{p_k})^T$ is also a solution of $F(z) = 0$; (ii) $\frac{\partial F}{\partial z}$ is invertible $\forall z\ (\in R^k)\ > 0$; and (iii) $z^* > 0$. By Lemma 1.6, $z = z^*$ must be an isolated point satisfying $F(z) = 0$. Therefore, there exists a $\delta_2 > 0$ ($\delta_2 \leq \delta_1$) such that for any $x \in \Omega_1(x^0) \cap U(x^*, \delta_2)$, $z = (x_{p_1}, \ldots, x_{p_r},\ x_{p_{r+1}}, \ldots, x_{p_k})^T$ is a solution of system (2.29) if and only if $x = x^*$, thus there is only one point $x^* \in \Omega_1(x^0) \cap U(x^*, \delta_2)$, i.e., x^* is an isolated point of $\Omega_1(x^0)$. But $\Omega_1(x^0)$ is connected, thus there is only one point x^* in $\Omega_1(x^0)$, the proof is complete. □

Theorem 2.9 ensures the strong convergence of the solution $x(t)$ of ODE (2.9) as $t \to +\infty$. This along with Theorem 2.7 guarantees the limit point is an optimal solution for problem (P_2). It should be mentioned that the limit point depends on the starting point x^0 in general.

2.3 An Affine Scaling Continuous Method for Box Constrained CQP

Based on the active set strategies, some algorithms for solving box constrained problems were presented [26, 32]. By solving a series of equality constrained quadratic optimization problems, finally the optimal solution of the original problem is obtained. But for large scale problems, there are two main disadvantages, one is that some constraints are added (dropped) at a time to (from) the working set, which leading to an excessive number of the iterations. The other disadvantage is that the exact minimizer on the current working face is required before adding (dropping) constraints [27]. In order to avoid these disadvantages, some gradient projection based algorithms were proposed [5, 41, 53, 54, 57, 77]. An algorithm that combines active set strategy with the gradient projection method was presented in [27]. Xia and Wang presented a projected dynamic system to solve the convex programming with box constraints. In order to ensure the convergence, where the strict convexity of objective function is required [77].

The presentation of this section is similar to the one in Section 2.2 but in a more compact format. In particular, many similar proofs will be omitted.

The KKT system of problem (P_3) is

$$\begin{cases} Qx + c = z - y, \ 0 \le x \le e, \\ (I_n - X)y = 0, \ y \ge 0, \\ Xz = 0, \ z \ge 0, \end{cases} \tag{2.30}$$

where $y, z \in R^n$.

In ODE system (2.9), X plays the role of a barrier wall such that the whole solution trajectory stays in the nonnegative region. Following the same idea, for problem (P_3), we consider the following ODE, which shares the similar properties of

ODE (2.9)

$$\begin{cases} \frac{dx(t)}{dt} = -X(I_n - X)(Qx + c), \\ x(0) = x^0, \ 0 < x^0 < e. \end{cases} \tag{2.31}$$

Theorem 2.10. *Let $x(t)$ be the solution of the system (2.31) with the maximal existence interval $[0,\ \beta)$. Then $0 < x(t) < e$ for any $t \in [0,\ \beta)$.*

Proof. We will prove that $0 < x(t) < e$ for any $t \in [0,\ \beta)$ by contradiction. In other words, $rank(X(I_n - X)) \equiv n$ for any $t \in [0,\ \beta)$.

Suppose that there exists a $t^* \in [0,\ \beta)$ such that $rank(X^*(I_n - X^*)) \leq n-1$. Since $x_i(t)$ is continuous on t, let t^* be the minimum t such that $rank(X^*(I_n - X^*)) \leq n-1$, i.e., $0 < x(t) < e$ for all $0 \leq t < t^*$. Thus there at least exists some $j \in \{1, \ldots, n\}$ such that $x_i(t^*) = 1$ or $x_i(t^*) = 0$. First suppose $x_i(t^*) = 1$, and

$$rank(X(t)(I_n - X(t))) = n,\ \forall t \in [0, t^*).$$

Let

$$M = \sup\{\|X(Qx + c)\| + 1 : 0 \leq x \leq e\}, \tag{2.32}$$

and

$$t_1 = \max\{0, t^* - \frac{1}{2M}\}.$$

Further, let $\bar{t}$ be the time satisfying

$$x_i(\bar{t}) = \min_{t \in [t_1, t^*]} x_i(t) < 1.$$

Notice that

$$\frac{dx(t)}{dt} = -(I_n - X)X(Qx + c),$$

we have

$$x_i(t^*) - x_i(t) = -\int_t^{t^*} (1 - x_i(\tau))e_i^T X(Qx + c)d\tau.$$

For any $t \in [t_1, t^*]$, $0 < x_i(t) < 1$, since $x_i(t^*) = 1$, we have

$$\begin{aligned} 1 - x_i(t) &\le M|t^* - t|(1 - x_i(\bar{t})) \\ &\le \frac{1}{2}(1 - x_i(\bar{t})). \end{aligned}$$

Since t is arbitrary in $[t_1, t^*]$, taking $t = \bar{t}$ in the above inequality, we have

$$1 - x_i(\bar{t}) \le \frac{1}{2}(1 - x_i(\bar{t})).$$

Thus, $x_i(\bar{t}) = 1$, which is a contradiction with the definition of $x_i(\bar{t})$, $x_i(t^*) = 1$ is rejected. Similarly, $x_i(t^*) = 0$ is rejected for some i and some $t^* \in [0,\ \beta)$. □

Thus, the matrix $X(I_n - X)$ plays a role of a box barrier such that the solution trajectory of the system (2.31) will stay in the box constrained region forever. By Lemma 1.1, $\beta = +\infty$.

Corollary 2.3. *There exists a unique solution $x(t)$ for ODE (2.31) on $[0,\ +\infty)$, and $0 < x(t) < e$ for any $t \in [0,\ +\infty)$.*

Theorem 2.11. *Let $x(t)$ be a solution of the system* (2.31) *on $[0,\ +\infty)$. Then* $\lim_{t\to+\infty} X(I_n - X)(Qx + c) = 0.$

Theorem 2.12. *(i) If x is an optimal solution of problem (P_3) and $0 < x < e$, then $X(I_n - X)(Qx + c) = 0$.(ii) Let $x(t)$ be the solution of the system* (2.31)*, If $X(I_n - X)(Qx + c)|_{x=x^0} = 0$, then $x(t) \equiv x^0$ for all $t \ge 0$. Moreover, x^0 is an optimal solution of the problem (P_3).*

Theorem 2.13. *Let $x(t)$ be the solution of ODE* (2.31)*, if $X(I_n - X)(Qx + c)|_{t=0} \neq 0$, then $X(I_n - X)(Qx + c) \neq 0$ for any $t \ge 0$.*

Similar to (2.14), define

$$u(x) = Qx + c. \tag{2.33}$$

Let $x(t)$ be the solution of the system (2.31). Define limit set

$$\Omega_2(x^0) = \{y \in R^n \mid \text{ is a cluster point of } x(t) \text{ of ODE (2.31)}\}. \tag{2.34}$$

Theorem 2.3 implies $0 < x(t) < e$ for any $t \geq 0$, thus the limit set $\Omega_2(x^0)$ is nonempty, compact, connected [15]. For some given $\bar{x} \in \Omega_2(x^0)$, let

$$\bar{u} = Q\bar{x} + c. \tag{2.35}$$

By the monotonicity of $q(x(t))$ and Theorem 2.11, we have the following corollary.

Corollary 2.4. *(i)* $q(x) = q(\bar{x})\ \forall x \in \Omega_2(x^0)$*; (ii)* $X(I_n - X)u(x) = 0\ \forall x \in \Omega_2(x^0)$*, where* $u(x)$ *is defined in (2.33).*

Let $\bar{J} = \{j | \bar{u}_j = 0,\ j \in \{1, \ldots, n\}\}$ and $\bar{J}^c = \{1, \ldots, n\} \setminus \bar{J}$. Then

$$\bar{u}_j \neq 0, \quad \text{for any } j \in \bar{J}^c.$$

Together with the equality $\bar{X}(I_n - \bar{X})\bar{u} = 0$, further we partition $\bar{J}^c$ by

$$\bar{J}^c_l = \{j | \bar{x}_j = 0,\ j \in \bar{J}^c\},\ \bar{J}^c_u = \{j | \bar{x}_j = 1,\ j \in \bar{J}^c\}. \tag{2.36}$$

Define set

$$\bar{\Lambda}_2 = \{x \in R^n \mid 0 \leq x \leq e,\ x_{\bar{J}^c_l} = 0,\ x_{\bar{J}^c_u} = 1,\ q(x) = q(\bar{x})\}, \tag{2.37}$$

clearly $\bar{\Lambda}_2$ is closed. $\bar{\Lambda}_2$ is nonempty since $\bar{x} \in \bar{\Lambda}_2$. The following results are similar to those in Section 2.2.

Theorem 2.14. $\bar{\Lambda}_2$ *is convex.*

Proof. Let x be an arbitrary point in the convex hull $co(\bar{\Lambda}_2)$, i.e., x is a positive linear convex combination of some points in $\bar{\Lambda}_2$, thus $x_{\bar{J}^c_l}$=0, $x_{\bar{J}^c_u}$=1, $0 \leq x \leq e$. Based on the convexity of $q(x)$, the following inequality holds

$$q(x) \leq q(\bar{x}). \tag{2.38}$$

On the other hand, let $\Delta x = x - \bar{x}$, then $(\Delta x)_{\bar{J}^c} = 0$, $\bar{u}^T(\Delta x) = 0$, again by the convexity of $q(x)$, we have

$$\begin{aligned} q(x) &\geq q(\bar{x}) + \nabla q(\bar{x})^T(\Delta x) \\ &= q(\bar{x}) + \bar{u}^T(\Delta x) \\ &= q(\bar{x}), \end{aligned}$$

this and (2.38) imply $q(x) = q(\bar{x})$ for all $x \in co(\bar{\Lambda}_2)$, thus $x \in \bar{\Lambda}_2$, hence $\bar{\Lambda}_2$ in convex. □

By Lemma 1.4 and Theorem 2.14, the following theorem is straightforward.

Theorem 2.15. $u(x) = \bar{u}$ *for all* $x \in \bar{\Lambda}_2$.

Theorem 2.16. $\Omega_2(x^0) \subseteq \bar{\Lambda}_2$.

Proof. If $\bar{J}_0^c$ is empty, $\bar{\Lambda}_2 = \{x \in R^n \mid 0 \leq x \leq e,\ q(x) = q(\bar{x})\}$. From Corollary 2.4 (i), the result holds clearly. Now we consider the case that $\bar{J}_0^c$ is nonempty. Suppose there exists a point $\hat{x} \in \Omega_2(x^0)$ but $\hat{x} \notin \bar{\Lambda}_2$. Notice that $\bar{\Lambda}_2$ lies inside the bounded set $\{x \in R^n \mid 0 \leq x \leq e\}$, so $\bar{\Lambda}_2$ is compact (since $\bar{\Lambda}_2$ is closed). Thus $u(x)$ in (2.33) is uniformly continuous over $\bar{\Lambda}_2$. Theorem 2.15 implies there exists some $\delta_0 > 0$ such that

$$|u_j(x)| \geq |\bar{u}_j|/2 > 0 \ \forall\ j \in \bar{J}_0^c,\ \forall\ x \in U(\bar{\Lambda}_2, \delta_0), \tag{2.39}$$

where $U(\bar{\Lambda}_2, \delta_0)$ is the δ_0-neighborhood of set $\bar{\Lambda}_2$. Since $\hat{x} \notin \bar{\Lambda}_2$ and $\bar{\Lambda}_2$ is compact, there exists some $\delta_1 \in (0,\ \delta_0] \cap (0,\ 0.1]$ such that $\hat{x} \notin U(\bar{\Lambda}_2, \delta_1)$. Notice $\bar{x} \in \Omega_2(x^0) \cap \bar{\Lambda}_2$ and $\hat{x} \in \Omega_2(x^0)$ but $\hat{x} \notin U(\bar{\Lambda}_2, \delta_1)$, by the connectivity of $\Omega_2(x^0)$, there must exist some $\tilde{x} \in \Omega_2(x^0) \cap U(\bar{\Lambda}_2, \delta_1)$ but $\tilde{x} \notin \bar{\Lambda}_2$. $\tilde{x} \in \Omega_2(x^0)$ and Corollary 2.4 (i) imply

$$0 \leq \tilde{x} \leq e,\ q(\tilde{x}) = q(\bar{x}).$$

Since $\tilde{x} \notin \bar{\Lambda}_2$, $\tilde{x} \in \Omega_2(x^0)$, and at least one of the sets $\bar{J}_l^c$ and $\bar{J}_u^c$ is nonempty, then at least one of the following two cases will occur

(a) $\tilde{x}_{\bar{J}_l^c} = 0$ is not true, i.e. there exists some $j_1 \in \bar{J}_l^c$ such that $\tilde{x}_{j_1} > 0$; or

(b) $\tilde{x}_{\bar{J}_u^c} = e_{\bar{J}_u^c}$ is not true, i.e. there exists some $j_2 \in \bar{J}_u^c$ such that $\tilde{x}_{j_2} < 1$.

If case (a) arises, since $\tilde{x} \in \Omega_2(x^0) \cap U(\bar{\Lambda}_2, \delta_1)$ and $\delta_1 \leq 0.1$, then $\tilde{x}_{j_1} < 0.1$. Thus, $0 < \tilde{x}_{j_1} < 0.1$.

If case (b) arises, since $\tilde{x} \in \Omega_2(x^0) \cap U(\bar{\Lambda}_2, \delta_1)$ and $\delta_1 \leq 0.1$, then $\tilde{x}_{j_2} > 0.9$. Thus, $0.9 < \tilde{x}_{j_2} < 1$.

In either case, there exists a j (j_1 or j_2) $\in \bar{J}_0^c$ such that $\tilde{x}_j(1-\tilde{x}_j) \neq 0$. (2.39) and $\delta_1 \leq \delta_0$ ensure $|u_j(\tilde{x})| > 0$, thus $\tilde{x}_j(1-\tilde{x}_j)u_j(\tilde{x}) \neq 0$. This contracts with the fact $\tilde{X}(I_n - \tilde{X})u(\tilde{x}) = 0$ (Corollary 2.4 (ii)) since $\tilde{x} \in \Omega_2(x^0)$, thus the proof is complete. □

Theorem 2.17. *If $x(t)$ is the solution of ODE (2.31), then $\lim_{t\to+\infty}(Qx+c) = \bar{u}$, and $\bar{u}_{\bar{J}_l^c} > 0$ if $\bar{J}_l^c$ is nonempty, $\bar{u}_{\bar{J}_u^c} < 0$ if $\bar{J}_u^c$ is nonempty.*

Proof. By the continuity of $s(x(t))$, compactness of the $\bar{\Lambda}_2$ and Theorem 2.16 and Theorem 2.17, clearly

$$\lim_{t\to+\infty}(Qx+c) = \bar{u}.$$

If $\bar{J}_l^c$ is nonempty, by the definition of $\bar{J}_l^c$, $\bar{u}_j \neq 0$ for any $j \in \bar{J}_l^c$. Suppose there exists some $\bar{j} \in \bar{J}_l^c$ such that $\bar{u}_{\bar{j}} < 0$. Since $u(x(t))$ is continuous on $[0, \infty)$, there exists some t_K such that $u_{\bar{j}}(x(t)) < 0$ for all $t \geq t_K$. For any cluster point $\bar{x} \in \Omega_2(x^0)$, $\bar{j} \in \bar{J}_l^c$ implies $\bar{x}_{\bar{j}} = 0$. Notice that

$$\frac{dx(t)}{dt} = -X(I_n - X)(Qx+c),$$

and $0 < x(t) < e$ for any $t \in [0, \infty)$. Thus $\frac{dx_{\bar{j}}(t)}{dt} \geq 0$, so $x_{\bar{j}}(t) \geq x_{\bar{j}}(t_K) > 0$ for all $t \geq t_K$, which contradicts with $\bar{x}_{\bar{j}} = 0$.

Similarly, we can prove $\bar{s}_{\bar{J}_u^c} < 0$ if $\bar{J}_u^c$ is nonempty. □

Theorem 2.18. *Any $x \in \Omega_2(x^0)$ is an optimal solution for problem (P_3).*

Proof. For any $x \in \Omega_2(x^0)$, we have

$$Qx + c = Q\bar{x} + c = \bar{u}$$

and

$$X(I_n - X)\bar{u} = 0.$$

By Theorem 2.17, together with the definition of $\bar{J}$, $\bar{J}_u^c$, $\bar{J}_l^c$, for any $i \in \{1, \ldots, n\}$, we have

$$\begin{aligned} \bar{u}_i &= 0, \quad \text{if } 0 < x_i < 1; \\ \bar{u}_i &\le 0, \quad \text{if } x_i = 1; \\ \bar{u}_i &\ge 0, \quad \text{if } x_i = 0. \end{aligned}$$

Let us define z, $y \in R^n$ by

$$\begin{cases} z_i = 0,\ y_i = 0, & \text{if } 0 < x_i < 1; \\ z_i = 0,\ y_i = -\bar{u}_i, & \text{if } x_i = 1; \\ z_i = \bar{u}_i,\ y_i = 0, & \text{if } x_i = 0. \end{cases} \tag{2.40}$$

Thus the following relations hold

$$\begin{cases} Qx + c = z - y,\ 0 \le x \le e, \\ (I_n - X)y = 0,\ y \ge 0, \\ Xz = 0,\ z \ge 0, \end{cases}$$

which are exactly the KKT conditions (2.30) of problem (P_3), thus the optimality of x is obvious. □

Theorem 2.19. *The limit set $\Omega_2(x^0)$ only contains a single point.*

Proof. From previous theorems in this section, $u(x) = Qx + c$ is a constant on $\Omega_2(x^0)$. Let u^* denote this constant. Assume $x^* \in \Omega_2(x^0)$, and the rank of matrix $X^*(I_n - X^*)$

is maximum for all $x \in \Omega_2(x^0)$, the index set $\{1, \ldots, n\}$ can be divided into three disjoint sets B, N_1, N_2 such that

$$0 < x_i^* < 1, \; i \in B,$$

$$x_i^* = 0, \qquad i \in N_1,$$

$$x_i^* = 1, \qquad i \in N_2,$$

further, let $N = N_1 \cup N_2$. If $B = \emptyset$, we can conclude there exist at most 2^n points in $\Omega_2(x^0)$, since $\Omega_2(x^0)$ is connected, there is only one point in $\Omega_2(x^0)$, the convergence of $x(t)$ is clear. So we suppose that B is nonempty, without loss of generality, let

$$B = \{1, \ldots, k\} \text{ and } N = \{k+1, \ldots, n\}.$$

Based on this partition, for any $x \in R^n$, it can be partitioned by $x = \begin{pmatrix} x_B \\ x_N \end{pmatrix}$. Similarly, we can partition $s = \begin{pmatrix} s_B \\ s_N \end{pmatrix}$, $c = \begin{pmatrix} c_B \\ c_N \end{pmatrix}$, $e = \begin{pmatrix} e_B \\ e_N \end{pmatrix}$ respectively, where x_B, s_B, c_B, $e_B \in R^k$, and x_N, s_N, c_N, $e_N \in R^{(n-k)}$.

Let $\delta_1 = \frac{1}{2}\min\{\min_{i\in B}\{x_i^*\}, \min_{i \in B}\{1 - x_i^*\}\}$. From the definition of x^*, we know

$$0 < x_B < e_B, \; x_N = x_N^*, \; \forall x \in \Omega_2(x^0) \cap U(x^*, \delta_1). \tag{2.41}$$

Next we will prove that x^* is an isolated point of $\Omega_2(x^0)$.

For any point $x \in \Omega_2(x^0) \cap U(x^*, \delta_1)$, the following equality holds

$$Qx + c = Qx^* + c \doteq u^*. \tag{2.42}$$

The above equality can be rewritten as

$$\begin{cases} Q_{11}x_1 + \cdots + Q_{1k}x_k + Q_{1(k+1)}x_{k+1} + \cdots + Q_{1n}x_n = u_1^* - c_1 \\ Q_{21}x_1 + \cdots + Q_{2k}x_k + Q_{2(k+1)}x_{k+1} + \cdots + Q_{2n}x_n = u_2^* - c_2 \\ \vdots \\ Q_{n1}x_1 + \cdots + Q_{nk}x_k + Q_{n(k+1)}x_{k+1} + \cdots + Q_{nn}x_n = u_n^* - c_n \end{cases} \tag{2.43}$$

For simplicity of our discussion, denote

$$Q = [q_1, q_2, \ldots, q_n],\ \ b = u^* - c - \sum_{i \in N} x_i^* q_i = u^* - c - \sum_{i \in N_2} q_i.$$

Since $(\Omega_2(x^0) \cap U(x^*, \delta_1)) \subseteq \Omega_2(x^0)$, we have

$$x_1 q_1 + \cdots + x_k q_k = b, \text{ for any } x \in \Omega_2(x^0) \cap U(x^*, \delta_1). \tag{2.44}$$

thus, $rank[q_1, q_2, \ldots, q_k] = rank[q_1, q_2, \ldots, q_k, b]$.

If $rank[q_1, q_2, \ldots, q_k] = k$, b will be expressed uniquely as a linear combination of $q_1, q_2, \ldots, q_k$, thus except for x^*, there is no x in $\Omega_2(x^0) \cap U(x^*, \delta)$ such that (2.44) holds. In other words, x^* is an isolated point of $\Omega_2(x^0)$.

If $rank[q_1, q_2, \ldots, q_k] = r < k$, the case that $r = 0$ is not under our consideration. If $r = 0$, then we have

$$\frac{dx_i(t)}{dt} = -c_i x_i (1 - x_i), \quad i = 1, \ldots, k.$$

From Theorem 2.11, $\lim_{t \to +\infty} \frac{dx_i(t)}{dt} = 0$, thus there are two cases

(a) $c_i = 0 \Rightarrow x_i(t) \equiv x_i^0$, thus the convergence is clear;

(b) $c_i \neq 0 \Rightarrow x_i^* = 0$ or $x_i^* = 1$, which is a contradiction with the assumption that $0 < x_i^* < 1$.

So we only consider the case that $1 \leq r < k$, and assume $\{q_{p_1}, q_{p_2}, \ldots, q_{p_r}\}$ is an maximum linearly independent group of $\{q_1, q_2, \ldots, q_k\}$, and $\{q_{p_{r+1}}, q_{p_{r+2}}, \ldots, q_{p_k}\} = \{q_1, q_2, \ldots, q_k\} \setminus \{q_{p_1}, q_{p_2}, \ldots, q_{p_r}\}$. Thus there exists a matrix $W = (w_{ij}) \in R^{(k-r) \times r}$ such that

$$q_{p_{r+i}} = \sum_{j=1}^{r} w_{ij} q_{p_j},\ \ i = 1, \ldots, k - r. \tag{2.45}$$

We consider the following sub-system of (2.43)

$$\begin{cases} q_{p_1}^T x &= (u^* - c)_{p_1} \\ \vdots & \\ q_{p_r}^T x &= (u^* - c)_{p_r} \\ q_{p_{r+1}}^T x &= (u^* - c)_{p_{r+1}} \\ \vdots & \\ q_{p_k}^T x &= (u^* - c)_{p_k} \end{cases} \tag{2.46}$$

Since x^* is a solution of (2.46), the system (2.46) is consistent. By (2.45), we have

$$(u^* - c)_{p_{r+i}} = \sum_{j=1}^{r} w_{ij} (u^* - c)_{p_j}, \ i = 1, \ldots, k - r. \tag{2.47}$$

$X^*(I_n - X^*)u^* = 0$ and $0 < x_B^* < e_B$ imply $u_B^* = 0$, thus

$$c_{p_{r+i}} = \sum_{j=1}^{r} w_{ij} c_{p_j}, \ i = 1, \ldots, k - r. \tag{2.48}$$

Let $x(t)$ be the solution of ODE (2.31). In order to overcome the difficulty caused by the degeneracy in linear equations (2.46), define

$$y_i(t) = \sum_{j=1}^{r} w_{ij} \ln \frac{x_{p_j}(t)}{1 - x_{p_j}(t)} - \ln \frac{x_{p_{r+i}}(t)}{1 - x_{p_{r+i}}}, \ i = 1, \ldots, k - r. \tag{2.49}$$

Since $0 < x(t) < e$ for any $t \in [0, +\infty)$, $y_i(t)$ is well defined for any $i \in \{1, \ldots, k-r\}$. Notice that

$$\frac{dx(t)}{dt} = -X(I_n - X)(Qx + c),$$

together with (2.45) and (2.48), we have

$$\begin{aligned} \frac{dy_i(t)}{dt} &= \sum_{j=1}^{r} w_{ij} \frac{\frac{dx_{p_j}(t)}{dt}}{x_{p_j}(1 - x_{p_j})} - \frac{\frac{dx_{p_{r+i}}(t)}{dt}}{x_{p_{r+i}}(1 - x_{p_{r+i}})} \\ &= (\sum_{j=1}^{r} w_{ij} q_{p_j}^T - q_{p_{r+i}}^T) x + \sum_{j=1}^{r} w_{ij} c_{p_j} - c_{p_{r+i}} \\ &\equiv 0, \quad i = 1, \ldots, k - r. \end{aligned}$$

Thus there exist $k - r$ constants $\bar{c}_i$ $(i = 1, \ldots, k - r)$ such that

$$y_i(t) \equiv \bar{c}_i, \ i = 1, \ldots, k - r, \ t \in [0, +\infty), \tag{2.50}$$

sequentially

$$\sum_{j=1}^{r} w_{ij} \ln \frac{x_{p_j}}{1-x_{p_j}} - \ln \frac{x_{p_{r+i}}}{1-x_{p_{r+i}}} \equiv \bar{c}_i,\ i=1,\dots,k-r,\ \forall x \in \Omega_2(x^0) \cap U(x^*,\delta_1). \quad (2.51)$$

Let $P \in R^{r\times k}$ be a matrix generated by choosing r linearly independent rows (say, $l_1,\dots,l_r$) from matrix $[q_{p_1},\dots,q_{p_r},q_{p_{r+1}},\dots,q_{p_k}]$. Then P can be written as

$$P = \begin{pmatrix} Q_{l_1p_1} & \cdots & Q_{l_1p_r} & Q_{l_1p_{r+1}} & \cdots & Q_{l_1p_k} \\ Q_{l_2p_1} & \cdots & Q_{l_2p_r} & Q_{l_2p_{r+1}} & \cdots & Q_{l_2p_k} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ Q_{l_rp_1} & \cdots & Q_{l_rp_r} & Q_{l_rp_{r+1}} & \cdots & Q_{l_rp_k} \end{pmatrix}$$

By the system (2.44), we have

$$P \begin{pmatrix} x_{p_1} \\ \vdots \\ x_{p_r} \\ x_{p_{r+1}} \\ \vdots \\ x_{p_k} \end{pmatrix} = \begin{pmatrix} b_{p_1} \\ \vdots \\ b_{p_r} \\ b_{p_{r+1}} \\ \vdots \\ b_{p_k} \end{pmatrix}, \forall x \in \Omega_2(x^0) \cap U(x^*,\delta_1), \quad (2.52)$$

where $L = \{l_1,\dots,l_r\}$. Combining system (2.51) and system (2.52), for any $x \in \Omega_2(x^0) \cap U(x^*,\delta_1)$, $z = (x_{p_1},\dots,x_{p_r},x_{p_{r+1}},\dots,x_{p_k})$ is a solution of the following nonlinear system

$$\begin{cases} Q_{l_1p_1}z_1 + \cdots + Q_{l_1p_r}z_r + Q_{l_1p_{r+1}}z_{r+1} + Q_{l_1p_{r+2}}z_{r+2} + \cdots + Q_{l_1p_k}z_k = b_{l1} \\ Q_{l_2p_1}z_1 + \cdots + Q_{l_2p_r}z_r + Q_{l_2p_{r+1}}z_{r+1} + Q_{l_2p_{r+2}}z_{r+2} + \cdots + Q_{l_2p_k}z_k = b_{l2} \\ \vdots \\ Q_{l_rp_1}z_1 + \cdots + Q_{l_rp_r}z_r + Q_{l_rp_{r+1}}z_{r+1} + Q_{l_rp_{r+2}}z_{r+2} + \cdots + Q_{l_rp_k}z_k = b_{lr} \\ w_{11} \ln \frac{z_1}{1-z_1} + \cdots + w_{1r} \ln \frac{z_r}{1-z_r} - \ln \frac{z_{r+1}}{1-z_{r+1}} - 0 - 0 - \cdots - 0 = \bar{c}_1 \\ w_{21} \ln \frac{z_1}{1-z_1} + \cdots + w_{2r} \ln \frac{z_r}{1-z_r} - 0 - \ln \frac{z_{r+2}}{1-z_{r+2}} - 0 - \cdots - 0 = \bar{c}_2 \\ \vdots \\ w_{(k-r)1} \ln \frac{z_1}{1-z_1} + \cdots + w_{(k-r)r} \ln \frac{z_r}{1-z_r} - 0 - 0 - 0 - \cdots - \ln \frac{z_k}{1-z_k} = \bar{c}_{k-r} \end{cases} \quad (2.53)$$

The Jacobian matrix of system (2.53) is

$$
J(z) = \begin{pmatrix}
Q_{l_1p_1} & \cdots & Q_{l_1p_r} & Q_{l_1p_{r+1}} & Q_{l_1p_{r+2}} & \cdots & Q_{l_1p_k} \\
Q_{l_2p_1} & \cdots & Q_{l_2p_r} & Q_{l_2p_{r+1}} & Q_{l_2p_{r+2}} & \cdots & Q_{l_2p_k} \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
Q_{l_rp_1} & \cdots & Q_{l_rp_r} & Q_{l_rp_{r+1}} & Q_{l_rp_{r+2}} & \cdots & Q_{l_rp_k} \\
\frac{w_{11}}{z_1(1-z_1)} & \cdots & \frac{w_{1r}}{z_r(1-z_r)} & -\frac{1}{z_{r+1}(1-z_{r+1})} & 0 & \cdots & 0 \\
\frac{w_{21}}{z_1(1-z_1)} & \cdots & \frac{w_{2r}}{z_r(1-z_r)} & 0 & -\frac{1}{z_{r+2}(1-z_{r+2})} & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{w_{(k-r)1}}{z_1(1-z_1)} & \cdots & \frac{w_{(k-r)r}}{z_r(1-z_r)} & 0 & 0 & \cdots & -\frac{1}{z_k(1-z_k)}
\end{pmatrix}.
$$

From (2.45), after a series of Gaussian operations, the Jacobian matrix $J(z)$ can be converted into

$$
\begin{aligned}
\bar{J}(z) &= \begin{pmatrix}
Q_{l_1p_1} & \cdots & Q_{l_1p_r} & 0 & 0 & \cdots & 0 \\
Q_{l_2p_1} & \cdots & Q_{l_2p_r} & 0 & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
Q_{l_rp_1} & \cdots & Q_{l_rp_r} & 0 & 0 & \cdots & 0 \\
\frac{w_{11}}{z_1(1-z_1)} & \cdots & \frac{w_{1r}}{z_r(1-z_r)} & \frac{-1}{z_{r+1}(1-z_{r+1})} - \sum_{j=1}^{r} \frac{w_{1j}^2}{z_j(1-z_j)} & -\sum_{j=1}^{r} \frac{w_{2j}w_{1j}}{z_j(1-z_j)} & \cdots & -\sum_{j=1}^{r} \frac{w_{(k-r)j}w_{1j}}{z_j(1-z_j)} \\
\frac{w_{21}}{z_1(1-z_1)} & \cdots & \frac{w_{2r}}{z_r(1-z_r)} & -\sum_{j=1}^{r} \frac{w_{1j}w_{2j}}{z_j(1-z_j)} & \frac{-1}{z_{r+2}(1-z_{r+2})} - \sum_{j=1}^{r} \frac{w_{2j}^2}{z_j(1-z_j)} & \cdots & -\sum_{j=1}^{r} \frac{w_{(k-r)j}w_{2j}}{z_j(1-z_j)} \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{w_{(k-r)1}}{z_1(1-z_1)} & \cdots & \frac{w_{(k-r)r}}{z_r(1-z_r)} & -\sum_{j=1}^{r} \frac{w_{1j}w_{(k-r)j}}{z_j(1-z_j)} & -\sum_{j=1}^{r} \frac{w_{2j}w_{(k-r)j}}{z_j(1-z_j)} & \cdots & \frac{-1}{z_k(1-z_k)} - \sum_{j=1}^{r} \frac{w_{(k-r)j}^2}{z_j(1-z_j)}
\end{pmatrix} \\
&\doteq \begin{pmatrix} Q_1 & 0 \\ Q_2 & Q_3 \end{pmatrix}.
\end{aligned}
$$

where $Q_1 \in R^{r\times r}$, $Q_2 \in R^{(k-r)\times r}$, and $Q_3 \in R^{(k-r)\times(k-r)}$. Based on the following two facts, $\bar{J}(z)$ (or $J(z)$) is invertible if $0 < z < e_B$,

(1) $rank(Q_1) = r$, and

(2)

$$Q_3 = -\begin{pmatrix} \frac{w_{11}}{\sqrt{z_1(1-z1)}} & \frac{w_{12}}{\sqrt{z_2(1-z_2)}} & \cdots & \frac{w_{1r}}{\sqrt{z_r(1-z_r)}} \\ \frac{w_{21}}{\sqrt{z_1(1-z_1)}} & \frac{w_{22}}{\sqrt{z_2(1-z_2)}} & \cdots & \frac{w_{2r}}{\sqrt{z_r(1-z_r)}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{w_{(k-r)1}}{\sqrt{z_1(1-z_1)}} & \frac{w_{(k-r)2}}{\sqrt{z_2(1-z_2)}} & \cdots & \frac{w_{(k-r)r}}{\sqrt{z_r(1-z_r)}} \end{pmatrix} \begin{pmatrix} \frac{w_{11}}{\sqrt{z_1(1-z_1)}} & \frac{w_{12}}{\sqrt{z_2(1-z_2)}} & \cdots & \frac{w_{1r}}{\sqrt{z_r(1-z_r)}} \\ \frac{w_{21}}{\sqrt{z_1(1-z_1)}} & \frac{w_{22}}{\sqrt{z_2(1-z_2)}} & \cdots & \frac{w_{2r}}{\sqrt{z_r(1-z_r)}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{w_{(k-r)1}}{\sqrt{z_1(1-z_1)}} & \frac{w_{(k-r)2}}{\sqrt{z_2(1-z_2)}} & \cdots & \frac{w_{(k-r)r}}{\sqrt{z_r(1-z_r)}} \end{pmatrix}^T$$

$$- \begin{pmatrix} \frac{1}{z_{r+1}(1-z_{r+1})} & 0 & \cdots & 0 \\ 0 & \frac{1}{z_{r+2}(1-z_{r+2})} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{z_k(1-z_k)} \end{pmatrix}.$$

Now let $G(z) = 0$ be system (2.53). From previous discussions, we know (i) $\forall x \in \Omega_2(x^0) \cap U(x^*, \delta_1)$, $z = (x_{p_1}, \ldots, x_{p_r}, x_{p_{r+1}}, \ldots, x_{p_k})^T$ is a solution of $G(z) = 0$, in particular, $z^* = (x^*_{p_1}, \ldots, x^*_{p_r}, x^*_{p_{r+1}}, \ldots, x^*_{p_k})^T$ is also a solution of $G(z) = 0$; (ii) $\frac{\partial G}{\partial z}$ is invertible $\forall z \in R^k$ and $0 < z < e_B$; and (iii) $0 < z^* < e_B$. By Lemma 1.6, $z = z^*$ must be an isolated point satisfying $G(z) = 0$. Therefore, there exists a $\delta_2 > 0$ ($\delta_2 \leq \delta_1$) such that for any $x \in \Omega_2(x^0) \cap U(x^*, \delta_2)$, $z = (x_{p_1}, \ldots, x_{p_r}, x_{p_{r+1}}, \ldots, x_{p_k})^T$ is a solution of system (2.53) if and only if $x = x^*$, thus there is only one point $x^* \in \Omega_2(x^0) \cap U(x^*, \delta_2)$, i.e., x^* is an isolated point of $\Omega_2(x^0)$. But $\Omega_2(x^0)$ is connected, thus there is only one point x^* in $\Omega_2(x^0)$, the proof is complete. □

2.4 Numerical Illustration

In order to illustrate the behaviors of ODEs (2.9) and (2.31), four small examples (two for each ODE) are constructed to depict the continuous trajectories. In addition, a set of ten randomly generated (Q, c)s are tested in Matlab with

$$\begin{cases} Q = AA^T,\ A = rand(n, r),\ r = 50 + round(rand(1) * (n - 50)); \\ c = \alpha * rand(n, 1) - e, \quad \alpha > 0 \text{ is a constant.} \end{cases}$$

The ODE solver used is **ODE23**. All of our tests were run in Matlab platform on a PC with 2.66 GHz processor. In our numerical tests, the conditional numbers of some Q's are very large, around 10^{20} for $n = 2000$.

2.4.1 The Performance of the Continuous Method for Nonnegativity Constrained CQP

Example 2.1.

$$\begin{array}{ll} \min & \frac{1}{2}(x_1^2 + 2x_1x_2 + 2x_1x_3 + x_2^2 + 2x_2x_3 + 3x_3^2) - \frac{1}{2}x_1 - x_2 + x_3 \\ s.t. & x_1, x_2, x_3 \geq 0. \end{array}$$

The unique optimal solution is $x^* = (0, 1, 0)^T$.

Example 2.2.

$$\begin{array}{ll} \min & \frac{1}{2}(x_1^2 + 2x_1x_2 + 2x_1x_3 + x_2^2 + 2x_2x_3 + 3x_3^2) - x_1 - x_2 + x_3 \\ s.t. & x_1, x_2, x_3 \geq 0. \end{array}$$

The optimal solution is $x^* = (\xi, 1 - \xi, 0)^T$, where $\xi \in [0,\ 1]$.

For three different random initial points, the solution trajectories for the above two examples are displayed in Figure 2.1. It is easy to see that all the trajectories will tend to some optimal solution, and the limit point depends on the initial point if there are more than one solutions.

The initial points for the ten randomly generated (Q, c)s are set to $x^0 = (1, \cdots, 1)^T$. The stopping criterion in our test is $|\frac{dx(t)}{dt}|_\infty \leq \epsilon$ for some small $\epsilon > 0$, which is guaranteed by Theorem 2.2 and $\frac{dq(x)}{dt} = -(Qx + c)^T X(Qx + c) \propto \left(|\frac{dx(t)}{dt}|_\infty\right)^2$. For $\epsilon = 10^{-4}$ and $\epsilon = 10^{-5}$, the numerical results with $\alpha = 5$ are reported in Table 2.1 and Table 2.2 respectively, where n represents the problem size (n) and CPU(s) represents the average time (run 100 times) required for solving each problem.

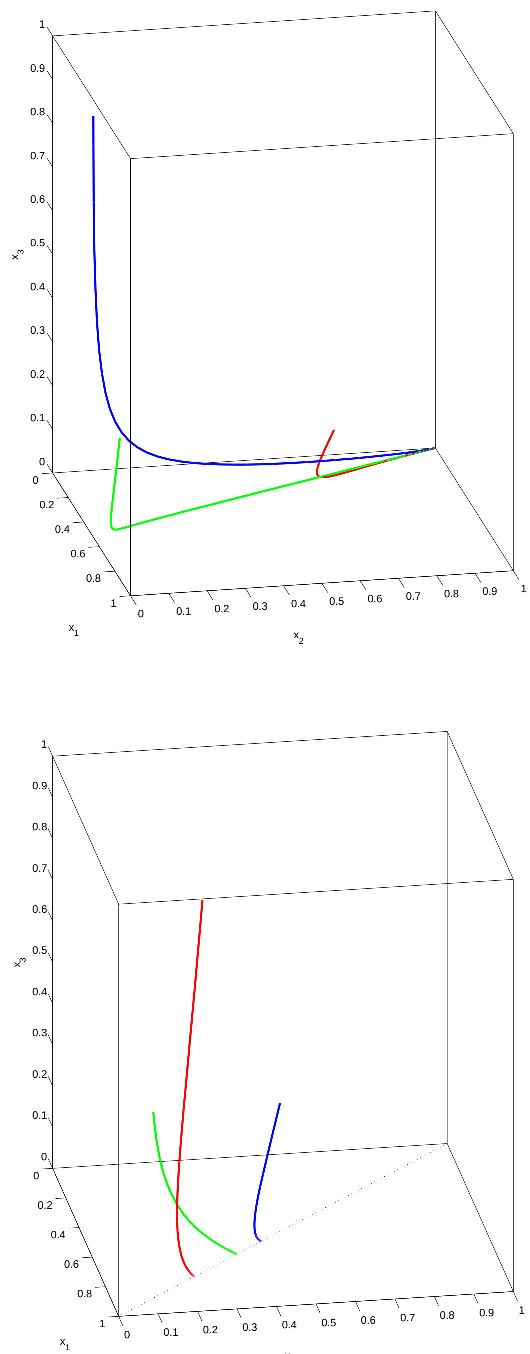

Figure 2.1: The limiting behavior of solution trajectories of ODE (2.9) starting from three different initial conditions. Upper: The solution trajectories for Example 2.1. Lower: The solution trajectories for Example 2.2.

Table 2.1: The performance of ODE (2.9) ($\epsilon = 10^{-4}$)

n	200	400	600	800	1000	1200	1400	1600	1800	2000
CPU(s)	0.02	0.03	0.09	0.23	0.33	0.44	0.51	0.67	0.97	1.01

Table 2.2: The performance of ODE (2.9) ($\epsilon = 10^{-5}$)

n	200	400	600	800	1000	1200	1400	1600	1800	2000
CPU(s)	0.04	0.08	0.19	0.51	0.80	0.90	0.92	1.14	1.26	1.61

2.4.2 The Performance of the Continuous Method for Box Constrained CQP

Example 2.3.

$$\begin{aligned} \min \quad & \tfrac{1}{2}x_1^2 + x_2^2 - 2x_1 - 3x_2 + 5x_3 \\ s.t. \quad & 0 \le x_1, x_2, x_3 \le 1. \end{aligned}$$

The unique optimal solution is $x^* = (1, 1, 0)^T$.

Example 2.4.

$$\begin{aligned} \min \quad & \tfrac{1}{2}(x_1^2 + 2x_1x_2 + 2x_1x_3 + x_2^2 + 2x_2x_3) - \tfrac{3}{2}x_1 - \tfrac{3}{2}x_2 + 5x_3 \\ s.t. \quad & 0 \le x_1, x_2, x_3 \le 1. \end{aligned}$$

The optimal solution is $x^* = (\xi, 1.5 - \xi, 0)^T$, where $\xi \in [0.5,\ 1]$.

For three different random initial points, the solution trajectories for the above two examples are displayed in Figure 2.2.

The initial points for the ten randomly generated (Q, c)s are set to $x^0 = (1/2, \cdots, 1/2)^T$. Similarly, for $\epsilon = 10^{-4}$ and $\epsilon = 10^{-5}$, the numerical results with $\alpha = 5$ are reported in Table 2.3 and Table 2.4 respectively, where n represents the problem size (n) and CPU(s) represents the average time (run 100 times) required for solving each problem.

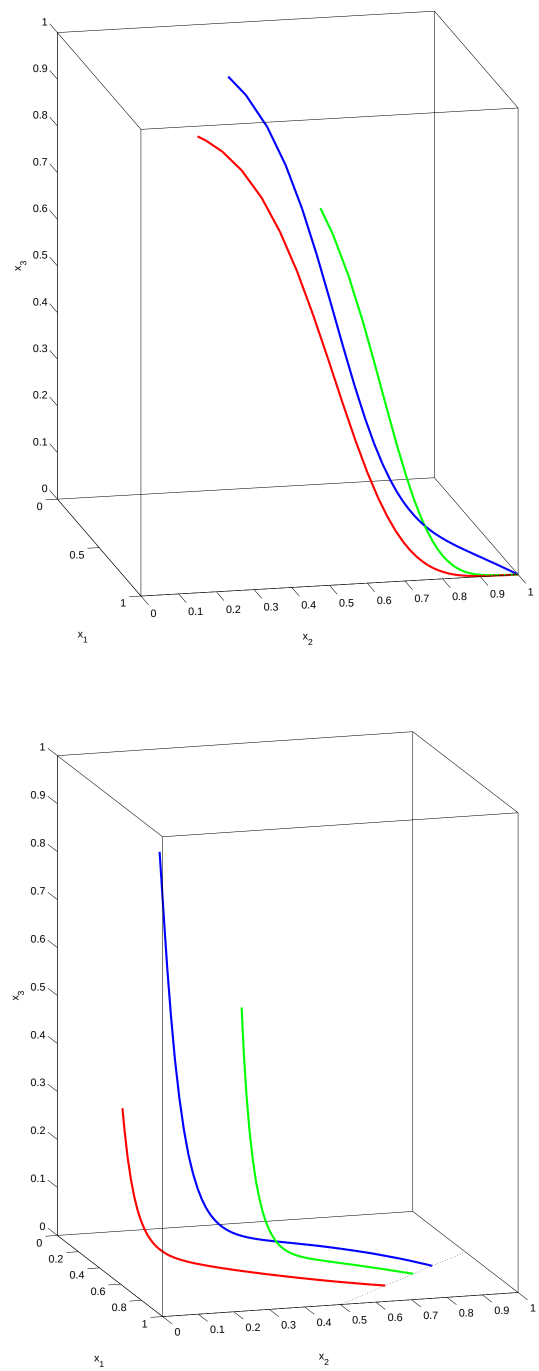

Figure 2.2: The limiting behavior of solution trajectories of ODE (2.31) starting from three different initial conditions. Upper: The solution trajectories for Example 2.3. Lower: The solution trajectories for Example 2.4.

Table 2.3: The performance of ODE (2.31) ($\epsilon = 10^{-4}$)

n	200	400	600	800	1000	1200	1400	1600	1800	2000
CPU(s)	0.02	0.04	0.08	0.24	0.28	0.50	0.57	0.64	0.81	1.00

Table 2.4: The performance of ODE (2.31) ($\epsilon = 10^{-5}$)

n	200	400	600	800	1000	1200	1400	1600	1800	2000
CPU(s)	0.05	0.09	0.19	0.53	0.61	0.83	1.04	0.84	1.41	1.97

From our preliminary results in Table 2.1, Table 2.2, Table 2.3 and Table 2.4, it can be observed that the performance of ODE (2.31) is more sensitive to the stopping criterion than that of ODE (2.9).

2.5 Concluding Remarks

Two interior point affine scaling continuous method models are introduced for non-negative and box constrained CQP respectively. The essence of each interior point affine scaling continuous method is an ODE system, whose equilibrium points correspond to the optimal solutions for the underlying optimization problem. A thorough study on the two continuous trajectories (one for each problem) is provided with many important theoretical results. In particular, strong convergence for both continuous trajectories is proved.

To extend the ODE (2.9) for problem (P_2), the following family of ODEs can be also used

$$\begin{cases} \frac{dx(t)}{dt} = -X^{\gamma}(Qx + c), \\ x(0) = x^0 > 0, \end{cases} \tag{2.54}$$

where $\gamma \geq 1$. Similar theoretical results obtained in Section 2.2 should also hold for ODE (2.54).

For general box constrained CQP (P_1) with $l < u$, if the optimal solution set is bounded, the following ODE system can be used

$$\begin{cases} \frac{dx(t)}{dt} = -g(x) \circ (Qx + c), \\ x(0) = x^0, \ l < x^0 < u, \end{cases} \tag{2.55}$$

where $\circ$ denotes the Hadamard product of two vectors, $g : R^n \to R^n$ is defined as follows

$$g_i(x) = \begin{cases} (u_i - x_i) & \text{if } l_i = -\infty, \ u_i < +\infty, \\ (u_i - x_i)(x_i - l_i) & \text{if } -\infty < l_i, \ u_i < +\infty, \\ (x_i - l_i) & \text{if } -\infty < l_i, \ u_i = +\infty, \\ 1 & \text{if } l_i = -\infty, \ u_i = +\infty, \end{cases}$$

$i = 1, \ldots, n$. Similarly, the theoretical results obtained in Section 2.3 should also hold for ODE (2.55).

Chapter 3

First-order Affine Scaling Continuous Trajectory for Standard CQP

3.1 Introduction

In this chapter we consider the following standard convex quadratic programming

$$\begin{aligned} \min \quad & q(x) = \frac{1}{2}x^T Qx + c^T x \\ s.t. \quad & Ax = b,\ x \geq 0, \end{aligned} \tag{P_4}$$

where $Q = (q_{ij})_{n\times n} \in R^{n\times n}$, $A = (a_{ij})_{m\times n} \in R^{m\times n}$, $c \in R^n$, $b \in R^m$. We also assume that Q is symmetric and positive semi-definite, in addition A has full row rank.

For later discussion, the following two regions are defined

$$\mathcal{P}^+ = \{x \in R^n | Ax = b,\ x \geq 0\},\ \mathcal{P}^{++} = \{x \in R^n | Ax = b,\ x > 0\},$$

where $\mathcal{P}^{++}$ is called the relative interior of $\mathcal{P}^+$. Since interior point method is considered, we assume $\mathcal{P}^{++}$ is nonempty.

For problem (P_4), the affine scaling algorithm was first proposed by Dikin in 1967. Due to its beautiful mathematical structure and high performance, many researchers turned to study the convergence of the affine scaling algorithm. For linear programming, please see [21, 59, 66, 68], for the continuous affine scaling trajectories, please see [4, 7, 45, 50] .

In Chapter 2, we introduced the idea of affine scaling algorithm briefly. Let's recall the first-order and second-order affine scaling algorithms for problem (P_4). Their search directions are obtained by minimizing a linear or quadratic function over the intersection of the null space of A with the ellipsoid centered at the current interior point [20, 33, 51, 63, 67, 82]. For some $x \in \mathcal{P}^{++}$, the linear and quadratic optimization subproblems can be stated as follows

$$\begin{aligned} \min_{d} \quad & (Qx+c)^T d \\ s.t. \quad & Ad = 0, \\ & \|X^{-1}d\|^2 \le \beta^2 < 1, \end{aligned} \tag{P_5}$$

$$\begin{aligned} \min_{d} \quad & (Qx+c)^T d + \tfrac{1}{2} d^T Q d \\ s.t. \quad & Ad = 0, \\ & \|X^{-1}d\|^2 \le \beta^2 < 1, \end{aligned} \tag{P_6}$$

where $\beta > 0$ is constant.

It is not hard to compute the optimal solutions of problems (P_5) and (P_6), i.e., the first-order and second-order affine scaling directions, which are expressed as follows

$$d_1 = -\frac{\beta X P_{AX} X (Qx+c)}{\|P_{AX} X (Qx+c)\|_2}, \tag{3.1}$$

$$d_2 = -\frac{\beta X P_{AX} X [Q(x+d_2)+c]}{\|P_{AX} X [Q(x+d_2)+c]\|_2}, \tag{3.2}$$

where $P_{AX} = I_n - XA^T(AX^2A^T)^{-1}AX$. It should be noted that formula (3.2) is an implicit scheme.

For the discrete methods, without the nondegeneracy assumption, Sun [63], Monteiro and Tsuchiya [51] proved the convergence of the second-order affine scaling algorithm. Compared with the second-order affine scaling algorithm, the structure of the first-order affine scaling algorithm is simpler, but the convergence has not been solved so far. Monteiro and Tsuchiya [51] pointed out the convergence of the first-order affine scaling algorithm had been proved by Gonzaga and Carlos in 1990. But in later version ([33] in 2002), Gonzaga and Carlos only gave weak convergence under primal nondegeneracy assumption and boundedness assumption of the optimal solution set. Under the same assumptions, Tseng, Bomze and Schachinger discussed the convergence of the generalized first-order affine scaling algorithm with Armijo-type step rule in [67], where the search direction is given by

$$d_\gamma = -X^\lambda P_{AX^\lambda} X^\lambda (Qx + c), \tag{3.3}$$

$\lambda > 0$ is constant. For any $\lambda \in (0,\ 1)$, they proved the convergence.

Especially, if $\lambda = \frac{1}{2}$, $A = (1, \ldots, 1)$, $b = 1$ and $c = 0$, the continuous algorithm induced by the direction d_γ reduces to the replicator dynamics. By constructing proper Lyapunov potential function, Losert and Akin proved the convergence of the continuous trajectory even though Q is not positive semi-definite [43].

In this chapter we consider the continuous trajectory determined by the following ODE system

$$\frac{dx}{dt} = -XP_{AX}X(Qx + c),\ x(0) = x^0 \in \mathcal{P}^{++}. \tag{3.4}$$

The rest of this chapter is organized as follows. In Section 3.2, to ensure that the ODE (3.4) is well defined, we give two standard assumptions, i.e., boundedness of the optimal solution set of problem (P_4) and nondegeneracy of the constrained region. Based on a special logarithmic barrier function optimization problem, we discuss the relationship between the centering direction and the affine scaling direction. For the affine scaling algorithm, the convergence proof methods for linear programming cannot be easily generalized to quadratic programming. In Section 3.3, a thorough study on the continuous trajectory determined by ODE system (3.4) will be investigated. In Section 3.4, with the help of dual estimate, the optimality of any accumulation point is proved. Moreover, the convergence of the dual estimate is obtained. In Section 3.5, the strong convergence of the continuous trajectory is proved. Finally, some concluding remarks are drawn in Section 3.6.

3.2 Central Path and Affine Scaling Direction

In this section, based on a logarithmic barrier function optimization problem, in the sense of continuity we discuss the relationship between the centering direction and the affine scaling direction briefly. First, we make the following two assumptions.

Assumption 3.1. *The optimal solution set of problem* (P_4) *is bounded.*

Assumption 3.2. *For any* $x \in \mathcal{P}^+$, AX^2A^T *is nonsingular, i.e., the columns of* A *corresponding to index set* $\{j \mid x_j \neq 0, j \in \{1, \ldots, n\}\}$ *have rank* m.

Assumption 3.2 is called the primal nondegeneracy assumption. From Lemma 1.3 (or Lemma 2.1 in [52]), the following lemma is straightforward.

Lemma 3.1. *Assumption 3.1 holds if and only if for any* $\bar{x} \in \mathcal{P}^+$, *the level set* $\{x \mid x \in \mathcal{P}^+,\ q(x) \leq q(\bar{x})\}$ *is bounded.*

Theorem 3.1. *Under Assumption 3.2,* $(AX^2A^T)^{-1} \in C^1$ *on* $\mathcal{P}^+$.

Proof. If Assumption 3.2 holds, then for any $x \in \mathcal{P}^+$, $(AX^2A^T)^{-1}$ exists, and $(AX^2A^T)^{-1}$ is continuous on $\mathcal{P}^+$. The following equality is straightforward

$$(AX^2A^T)(AX^2A^T)^{-1} = I_m. \tag{3.5}$$

For any $i \in \{1, \ldots, n\}$, taking the partial derivative with respect to x_i in the equality (3.5), we have

$$2x_i(Ae_ie_i^TA^T)(AX^2A^T)^{-1} + (AX^2A^T)\frac{\partial(AX^2A^T)^{-1}}{\partial x_i} = 0,$$

consequently

$$\frac{\partial(AX^2A^T)^{-1}}{\partial x_i} = -2x_i(AX^2A^T)^{-1}(Ae_ie_i^TA^T)(AX^2A^T)^{-1}. \tag{3.6}$$

Thus the proof is complete. □

Next we discuss the relationship between central path and affine scaling direction. Considering the following optimization problem

$$\begin{aligned} \min_{x} \quad & \tfrac{1}{2}x^T Qx + c^T x - \mu \sum_{i=1}^{n} \ln x_i \\ s.t. \quad & Ax = b,\ x > 0, \end{aligned} \tag{$P_4\mu$}$$

where $\mu > 0$ is the barrier penalty parameter. It is easy to verify that problem $(P_4\mu)$ has no more than one minimizer. For given μ, if the minimizer exists, denoted by $x(\mu)$, then it is determined by the following KKT system

$$\begin{cases} Ax = b,\ x > 0, \\ Xz = \mu e,\ z > 0, \\ Qx + c - z - A^T y = 0, \end{cases} \tag{3.7}$$

where $y \in R^m$, $z \in R^n$.

Let $(x(\mu), y(\mu), z(\mu))$ be the solution of the above system (3.7). If for any $\mu > 0$, $(x(\mu), y(\mu), z(\mu))$ exists, then we get one smooth trajectory, which is called the central path, denoted by

$$\{(x(\mu), y(\mu), z(\mu)) \mid \mu > 0\}. \tag{3.8}$$

Lemma 3.2. *[23] Under Assumption 3.1, the central path (3.8) is well defined, and $x(\mu)$ converges as $\mu \to 0$ to the analytic center of the optimal solution set of problem (P_4).*

For more theoretical results about the central path, please see Monteiro and Zhou [52], or Drummond and Svaiter [23]. Of course, it is very hard to get the explicit solution $(x(\mu), y(\mu), z(\mu))$, taking the derivative with respective to μ in (3.7)

$$\begin{cases} A\frac{dx}{d\mu} = 0, \\ X\frac{dz}{d\mu} + Z\frac{dx}{d\mu} = e, \\ Q\frac{dx}{d\mu} - \frac{dz}{d\mu} - A^T\frac{dy}{d\mu} = 0, \end{cases} \tag{3.9}$$

where $Z = diag(z_1, z_2, \ldots, z_n) \in R^{n\times n}$. After series of calculation, we have

$$\frac{dx}{d\mu} = \frac{1}{\mu^2} X P_{AX} X[Q(x - \mu\frac{dx}{d\mu}) + c]. \tag{3.10}$$

Let $t = \frac{1}{\mu}$. Then (3.10) becomes

$$\frac{dx}{dt} = -XP_{AX}X[Q(x + t\frac{dx}{dt}) + c]. \tag{3.11}$$

For linear programming, in the continuous version, (3.11) implies that the centering direction and the affine scaling direction are completely consistent. Furthermore, if the initial point x^0 falls on the central path, the affine scaling continuous trajectory is exactly the central path, and converges to the analytic center of the optimal solution set. If the initial point is not on the central path, the affine scaling continuous trajectory can be characterized as the solution of certain parametrized logarithmic barrier optimization problem, i.e., one term $-\mu p^T x$ is added to the objective function in problem $(P_4\mu)$, thus the convergence is also guaranteed [4].

For quadratic programming, the centering direction (3.11) looks like the second-order affine scaling direction, but it is quite different from the first-order affine scaling direction. It should be noted that $\frac{dx}{dt}$ in (3.11) will remain unchanged if $-\mu p^T x$ is added to the objective function in problem $(P_4\mu)$. So far, it has not been clear whether ODE (3.4) can be regarded as the solution of certain optimization problem as described in linear programming.

3.3 Properties of the Continuous Trajectory

In this section, under Assumption 3.1 and Assumption 3.2, we will adopt the framework of continuous methods outlined in [42] to analyze some properties of the ODE (3.4), especially the weak convergence. For given starting point x^0 in ODE (3.4), let's define level set

$$L_2(x^0) = \{x \in \mathcal{P}^+ \mid q(x) \leq q(x^0)\}. \tag{3.12}$$

It is easy to verify that $XP_{AX}X(Qx + c)$ is locally Lipschitz continuous on $\{x \in R^n | x > 0\}$. From Lemma 1.1, there exists a unique solution $x(t)$ of ODE (3.4) on the interval $[0,\ \beta)$, for some $\beta > 0$.

Theorem 3.2. *Let $x(t)$ be the solution of ODE (3.4) with the maximal existence interval* $[0,\ \beta)$. *Then $Ax(t) = b$, $\forall t \in [0,\ \beta)$.*

Proof. Notice that for any $t \in [0,\ \beta)$

$$x(t) = x^0 - \int_{t_0}^{t} (XP_{AX}X(Qx + c)|_{t=\tau})d\tau,$$

thus

$$Ax(t) = Ax^0 - \int_{t_0}^{t} (AXP_{AX}X(Qx + c)|_{t=\tau})d\tau = b.$$

□

Theorem 3.3. *Let $x(t)$ be the solution of ODE (3.4) with the maximal existence interval* $[0,\ \beta)$. *Then $x(t) > 0$, $\forall t \in [0,\ \beta)$.*

Proof. We will prove $x(t) > 0$ for any $t \in [0,\ \beta)$ by contradiction. Suppose that there exists a $t^* \in (0,\ \beta)$, and an $i \in \{1, \ldots, n\}$ such that $x_i(t^*) = 0$. Since $x_i(t)$ is continuous on t, let t^* be the minimum t such that $x_i(t) = 0$, or rather, $x_i(t^*) = 0$ and

$$x(t) > 0,\ \forall t \in [0,\ t^*). \tag{3.13}$$

Theorem 3.2 implies

$$Ax(t) = b,\ \forall t \in [0,\ t^*]. \tag{3.14}$$

Notice that

$$\frac{dq(x(t))}{dt} = \nabla q(x)^T \frac{dx}{dt} = -(Qx + c)^T XP_{AX}X(Qx + c) \leq 0,\ \forall t \in [0,\ t^*], \tag{3.15}$$

thus $q(x(t))$ is nonincreasing monotonically along the trajectory $x(t)$. (3.13) and (3.14) together with (3.15) imply

$$x(t) \in L_2(x^0),\ \forall t \in [0,\ t^*].$$

Since $A^T(AX^2A^T)^{-1}AX^2$ is continuously differentiable on $L_2(x^0)$, let

$$\bar{\chi} = \max_{x \in L_2(x^0)} \|[I_n - A^T(AX^2A^T)^{-1}AX^2](Qx+c)\| + 1, \tag{3.16}$$

and

$$\delta = \max_{t\in[0,\ t^*]} x_i(t) > 0, \quad t_1 = \max\{0, t^* - \frac{1}{2\bar{\chi}\delta}\}.$$

Furthermore, let $\bar{t}$ be the time satisfying

$$x_i(\bar{t}) = \max_{t\in[t_1,\ t^*]} x_i(t) > 0.$$

From ODE (3.4), for any $t \in [t_1,\ t^*]$, we have

$$x_i(t^*) - x_i(t) = -\int_t^{t^*} x_i^2(\tau) e_i^T([I_n - A^T(AX^2A^T)^{-1}AX^2](Qx+c)|_{t=\tau})d\tau. \tag{3.17}$$

It follows from (3.17) that

$$|x_i(t^*) - x_i(t)| \le \bar{\chi} x_i^2(\bar{t})(t^* - t).$$

Since $x_i(t^*) = 0$ and $x_i(t) \ge 0$ on $[t_1,\ t^*]$, for any $t \in [t_1,\ t^*]$, we have

$$x_i(t) \le \bar{\chi} x_i^2(\bar{t})(t^* - t_1) \le \delta\bar{\chi} x_i(\bar{t})(t^* - t_1).$$

Taking $t = \bar{t}$, then

$$x_i(\bar{t}) \le \delta\bar{\chi} x_i(\bar{t})(t^* - t_1) \le \frac{1}{2} x_i(\bar{t}),$$

which is a contradiction with $x_i(\bar{t}) > 0$. $\square$

If $[0,\ \beta)$ is the maximal existence interval of the solution $x(t)$ of ODE (3.4), together with (3.15), Theorem 3.2 and Theorem 3.3 imply the whole solution trajectory $x(t)$ is contained in the compact set $L_2(x^0)$. By the continuation theorem (Lemma 1.1), the following theorem is straightforward.

Theorem 3.4. *There exists a unique solution $x(t)$ of ODE (3.4) on $[0,\ +\infty)$, and $x(t) \in \mathcal{P}^{++}$ for any $t \in [0,\ +\infty)$.*

Theorem 3.5. *For some $x \in \mathcal{P}^{++}$, x is an optimal solution for problem (P_4) if and only if $P_{AX}X(Qx + c) = 0$.*

Proof. The KKT conditions for problem (P_4) can be stated as follows

$$\begin{cases} Ax = b, \ \ x \geq 0, \\ Xz = 0, \ \ z \geq 0, \\ A^T y + z = Qx + c, \end{cases} \tag{3.18}$$

where $z \in R^n$, and $y \in R^m$.

If $x \in \mathcal{P}^{++}$ is an optimal solution, there exists corresponding (y, z) such that system (3.18) holds. Thus

$$\begin{cases} z = 0, \\ A^T y = Qx + c. \end{cases} \tag{3.19}$$

It is easy to check that

$$P_{AX}X(Qx + c) = P_{AX}XA^T y = 0.$$

The necessity is proved.

Conversely, if $x \in \mathcal{P}^{++}$ and

$$P_{AX}X(Qx + c) = X[I_n - A^T(AX^2A^T)^{-1}AX^2](Qx + c) = 0, \tag{3.20}$$

then

$$[I_n - A^T(AX^2A^T)^{-1}AX^2](Qx + c) = 0.$$

Let $\tilde{z} = 0 \in R^n$ and $\tilde{y} = (AX^2A^T)^{-1}AX^2(Qx+c)$. Then triple $(x, \tilde{y}, \tilde{z})$ satisfies KKT system (3.18). The sufficiency is also proved. □

From the point of view of ODE (or by the proofs of Theorem 2.3 and Theorem 2.4 in Chapter 2), the following two properties are trivial.

Proposition 3.1. *If $x(t)$ is the solution of ODE (3.4) and $XP_{AX}X(Qx+c)|_{t=0} \neq 0$, then $XP_{AX}X(Qx + c) \neq 0$ for any $t \geq 0$.*

Proposition 3.2. *If $x(t)$ is the solution of ODE (3.4) and $XP_{AX}X(Qx+c)|_{t=0} = 0$, then $XP_{AX}X(Qx+c) \equiv 0$ on $[0,\ +\infty)$.*

The following theorem implies the weak convergence of ODE system (3.4).

Theorem 3.6. *Let $x(t)$ be the solution of ODE (3.4). Then $\lim\limits_{t\to+\infty} XP_{AX}X(Qx+c) = 0$.*

Proof. If $x(t)$ is the solution of ODE (3.4), then $x(t)$ is contained in compact set $L_2(x^0)$. Thus (3.15) implies $q(x)$ has a finite limit as $t \to +\infty$. Obviously, $(Qx + c)^T XP_{AX}X(Qx + c)$ is continuously differentiable with respect to x on compact set $L_2(x^0)$. Since $x(t)$ is bounded on $L_2(x^0)$, there exists a constant $K_1 > 0$ such that

$$\begin{aligned} \left|\frac{dq(x)}{dt}\right|_{t=t_1} - \frac{dq(x)}{dt}\Big|_{t=t_2}\Big| &\le K_1\|x(t_1) - x(t_2)\| \\ &= K_1\left\|\int_{t_1}^{t_2} XP_{AX}X(Qx+c)dt\right\| \\ &\le K_1K_2|t_1 - t_2|, \end{aligned}$$

where $K_2 = \max\limits_{x\in L_2(x^0)} \|XP_{AX}X(Qx + c)\|$. Thus $\frac{dq(x)}{dt}$ is uniformly continuous on $[0,\ +\infty)$, Barbalat's Lemma 1.2 ensures that

$$\lim_{t\to+\infty} (Qx + c)^T XP_{AX}X(Qx + c) = 0.$$

Clearly $XP_{AX}X$ is positive semi-definite and bounded on $L_2(x^0)$, thus

$$\lim_{t\to+\infty} XP_{AX}X(Qx + c) = 0.$$

□

3.4 Optimality of Cluster Point

In this section, we discuss the optimality of any accumulation point of $x(t)$, which is the solution of the ODE (3.4), and show that the dual estimate is convergent. The proofs of main results in this section are similar to those in [67].

For any $x \in \mathcal{P}^+$, define dual estimate $(y(x), z(x))$ as follows

$$y(x) = (AX^2A^T)^{-1}AX^2(Qx + c), \tag{3.21}$$

$$z(x) = [I_n - A^T(AX^2A^T)^{-1}AX^2](Qx + c). \tag{3.22}$$

We have

$$\begin{cases} A^Ty(x) + z(x) = Qx + c, \\ Ax = b, \;\; x \geq 0. \end{cases} \tag{3.23}$$

Note: If $x(t)$ is the solution of ODE (3.4), $z(x(t))$ may not be nonnegative for any $t \in [0, \; +\infty)$.

Let $x(t)$ be the solution of the ODE (3.4). Then $x(t) \in L_2(x^0)$ for any $t \in [0, \; +\infty)$. Let's define limit set

$$\Omega_3(x^0) = \{p \in \mathcal{P}^+ \mid \; p \text{ is a cluster point of } x(t) \text{ of ODE (3.4)}\}.$$

$L_2(x^0)$ is bounded. Thus $\Omega_3(x^0)$ is nonempty, connected, and compact [15].

For some $\bar{x} \in \Omega_3(x^0)$, define $(\bar{z}, \bar{y})$ as follows

$$\bar{z} = z(\bar{x}) = [I_n - A^T(A\bar{X}^2A^T)^{-1}A\bar{X}^2](Q\bar{x} + c), \tag{3.24}$$

and

$$\bar{y} = y(\bar{x}) = (A\bar{X}^2A^T)^{-1}A\bar{X}^2(Q\bar{x} + c). \tag{3.25}$$

Clearly

$$Q\bar{x} + c = \bar{z} + A^T\bar{y}. \tag{3.26}$$

(3.15) and Theorem 3.6 imply the following proposition.

Proposition 3.3. *If $x \in \Omega_3(x^0)$, then*

(i) $q(x) = q(\bar{x})$;

(ii) $Xz(x) = 0$, where $z(x)$ is defined in (3.22).

Let $\bar{J} = \{j | \bar{z}_j = 0,\ j \in \{1, \ldots, n\}\}$, $\bar{J}^c = \{1, \ldots, n\} \setminus \bar{J}$. Define

$$\bar{\Lambda}_3 = \{x \in \mathcal{P}^+ \mid \ x_{\bar{J}^c} = 0,\ q(x) = q(\bar{x})\}, \tag{3.27}$$

proposition 3.3 (ii) implies $\bar{x}_{\bar{J}^c} = 0$, so $\bar{x} \in \bar{\Lambda}_3$. Thereby $\bar{\Lambda}_3$ is nonempty. In addition, it is easy to verify that $\bar{\Lambda}_3$ is closed.

Theorem 3.7. $\bar{\Lambda}_3$ *is convex.*

Proof. For any two points $\tilde{x}$ and $\check{x}$ in $\bar{\Lambda}_3$, and any $\lambda \in [0, 1]$, let $\hat{x} = \lambda\tilde{x} + (1 - \lambda)\check{x}$. Then

$$\hat{x}_{\bar{J}^c} = 0,\ A\hat{x} = b,\ \hat{x} \geq 0.$$

From Proposition 3.3 (i), along with the convexity of $q(x)$, we have

$$q(\hat{x}) \leq \lambda q(\tilde{x}) + (1 - \lambda)q(\check{x}) = q(\bar{x}).$$

On the other hand, let $\Delta x = \hat{x} - \bar{x}$. Then $(\Delta x)_{\bar{J}^c} = 0$, $A\Delta x = 0$ and $\bar{z}^T \Delta x = 0$. From (3.26) and the convexity of $q(x)$, we have

$$\begin{aligned} q(\hat{x}) &\geq q(\bar{x}) + \nabla q(\bar{x})^T(\Delta x) \\ &= q(\bar{x}) + (\bar{z} + A^T\bar{y})^T(\Delta x) \\ &= q(\bar{x}). \end{aligned}$$

So $q(\hat{x}) = q(\bar{x})$, thus $\hat{x} \in \bar{\Lambda}_3$. This completes the proof. □

Theorem 3.8. $z(x) = \bar{z}$ *for all* $x \in \bar{\Lambda}_3$.

Proof. For any $x \in \bar{\Lambda}_3$, (3.27) implies $X\bar{z} = 0$. Theorem 3.7 and Lemma 1.4, together with the definition of $\bar{\Lambda}_3$ and (3.26) ensure $Qx + c = Q\bar{x} + c = \bar{z} + A^T\bar{y}$. Thus

$$\begin{aligned} z(x) &= [I_n - A^T(AX^2A^T)^{-1}AX^2](Qx + c) \\ &= [I_n - A^T(AX^2A^T)^{-1}AX^2](\bar{z} + A^T\bar{y}) \\ &= [I_n - A^T(AX^2A^T)^{-1}AX^2]\bar{z} \\ &= \bar{z}. \end{aligned}$$

$\square$

Theorem 3.9. $\Omega_3(x^0) \subseteq \bar{\Lambda}_3$.

Proof. If $\bar{J}^c$ is empty, then $\bar{\Lambda}_3 = \{x \in \mathcal{P}^+ | \ q(x) = q(\bar{x})\}$. From Proposition 3.3 (i), the result holds clearly. If $\bar{J}^c$ is nonempty, suppose there is one point $\hat{x} \in \Omega_3(x^0)$ but $\hat{x} \notin \bar{\Lambda}_3$. Then $\hat{x} \in \mathcal{P}^+$, $q(\hat{x}) = q(\bar{x})$, but $\hat{x}_{\hat{j}} > 0$ for some $\hat{j} \in \bar{J}^c$. Clearly $\bar{\Lambda}_3$ lies inside the bounded level set $L_2(x^0)$, so $\bar{\Lambda}_3$ is compact. Thus $z(x)$ is uniformly continuous over $\bar{\Lambda}_3$. There exists some $\delta_0 > 0$ such that

$$|z_j(x)| \geq |\bar{z}_j|/2 \ \forall j \in \bar{J}^c, \ \forall x \in U(\bar{\Lambda}_3, \delta_0) \cap \mathcal{P}^+, \tag{3.28}$$

where $U(\bar{\Lambda}_3, \delta_0)$ is the δ_0-neighborhood of set $\bar{\Lambda}_3$. Let $\delta = \min\{\delta_0, \frac{\hat{x}_{\hat{j}}}{2}\}$. Then $\hat{x} \notin U(\bar{\Lambda}_3, \delta)$. Since $\Omega_3(x^0)$ is connected, $\bar{x} \in \Omega_3(x^0)$, $\hat{x} \in \Omega_3(x^0)$ and

$$\bar{x} \in \bar{\Lambda}_3, \ \hat{x} \notin U(\bar{\Lambda}_3, \delta),$$

there must exist some point $\tilde{x} \in \Omega_3(x^0)$ such that

$$\tilde{x} \in U(\bar{\Lambda}_3, \delta), \ \tilde{x} \notin \bar{\Lambda}_3. \tag{3.29}$$

In addition, $\tilde{x} \in \Omega_3(x^0)$ implies

$$\tilde{x} \in \mathcal{P}^+, \ q(\tilde{x}) = q(\bar{x}). \tag{3.30}$$

(3.29) and (3.30) imply there exists some $r \in \bar{J}^c$ such that $\tilde{x}_r \neq 0$. Since $\delta \leq \delta_0$, (3.28) implies $|z_j(\tilde{x})| \geq |\bar{z}_j|/2 > 0$ for all $j \in \bar{J}^c$. Thus $\tilde{x}_r z_r(\tilde{x}) \neq 0$, which contradicts with the fact $\tilde{X} z(\tilde{x}) = 0$ (Proposition 3.3 (ii)) since $\tilde{x} \in \Omega_3(x^0)$. $\square$

Based on the continuity of $z(x(t))$, compactness of the $\bar{\Lambda}_3$, together with Theorem (3.8) and Theorem (3.9), the following theorem is straightforward.

Theorem 3.10. *If $x(t)$ is the solution of ODE* (3.4), *then* $\lim_{t \to +\infty} z(x(t)) = \bar{z}$.

Since A has full row rank, form (3.23) and Theorem 3.10, the following theorem is also obvious.

Theorem 3.11. *If $x(t)$ is the solution of ODE* (3.4)*, then* $\lim_{t \to +\infty} y(x(t)) = \bar{y}$.

Similar to the proof of Theorem 2.6, the following theorem is clear.

Theorem 3.12. $\bar{z} \geq 0$.

Theorem 3.13. *Any point $x \in \Omega_3(x^0)$ is an optimal solution of problem (P_4).*

Proof. For any $x \in \Omega_3(x^0)$, $Ax = b$, $x \geq 0$. Proposition 3.3 (ii) and Theorem 3.10 imply

$$Xz(x) = X\bar{z} = 0.$$

From Theorem 3.8 and Theorem 3.9, we know

$$Qx + c = Q\bar{x} + c.$$

Together with (3.26), triple $(x, \bar{y}, \bar{z})$ satisfies the following system

$$\begin{cases} Ax = b, \ x \geq 0, \\ X\bar{z} = 0, \ \bar{z} \geq 0, \\ A^T\bar{y} + \bar{z} = Qx + c. \end{cases} \tag{3.31}$$

Thus the optimality of x is straightforward. □

3.5 Strong Convergence

For linear programming, there are two techniques can be used to prove the convergence of the affine scaling continuous trajectory. The first one has been described in Section 3.2, for more details, please see [4]. The other is the angle condition, which can be stated by the following lemma.

Lemma 3.3. *[66, 69] There exists a positive constant $\Delta(A, c)$ which is determined from A and c such that*

$$c^T X P_{AX} X c \geq \Delta(A, c)\|c\|\|X P_{AX} X c\|, \ \forall x \in \mathcal{P}^{++}.$$

Absil, Mahony and Andrews gave a sufficient condition to ensure the strong convergence of the general dynamical system, which is stated as follows.

Lemma 3.4. *[2] Let ϕ be a real analytic function [1] and let $x(t)$ be a C^1 curve in R^n, with $\dot{x}(t) = \frac{dx(t)}{dt}$ denoting its time derivative. Assume that there exist a $\delta > 0$ and a real τ such that for $t > \tau$, $x(t)$ satisfies the angle condition*

$$\frac{d\phi(x(t))}{dt} \equiv \langle \nabla\phi(x(t)), \dot{x}(t)\rangle \leq -\delta \cdot \|\nabla\phi(x(t))\| \cdot \|\dot{x}(t)\|,$$

and a weak decrease condition

$$[\frac{d}{dt}\phi(x(t)) = 0] \Rightarrow [\dot{x}(t) = 0].$$

Then, either $\lim_{t\to+\infty} \|x(t)\| = +\infty$ or there exists $x^* \in R^n$ such that $\lim_{t\to+\infty} x(t) = x^*$.

Lemma 3.3 together with Lemma 3.4 ensures the strong convergence of the affine scaling trajectories for linear programming. A natural question arises, does the angle condition hold for CQP? The following example indicates that the objective function $q(x)$ may not be a proper energy function used to prove the angle condition.

Example 3.1.

$$\begin{aligned} \min \quad & \tfrac{1}{2}(x_1^2 + x_2^2) + x_3 \\ s.t. \quad & x_1 = 1, \ x_1, x_2, x_3 \geq 0. \end{aligned}$$

If the ODE (3.4) is used to solve this problem, then

$$\frac{dx}{dt} = -XP_{AX}X(Qx + c) = -\begin{pmatrix} 0 \\ x_2^3 \\ x_3^2 \end{pmatrix}.$$

[1] A real function is said to be analytic if it possesses derivatives of all orders and agrees with its Taylor series in the neighborhood of every point.

Furthermore, x can be given explicitly by $x = (1, \frac{1}{\sqrt{2(t+k_1)}}, \frac{1}{t+k_2})^T$, where k_1, k_2 are two constants depend on the initial condition. If $\phi(x) = \frac{1}{2}(x_1^2 + x_2^2) + x_3$, then

$$\begin{aligned}
\lim_{t\to+\infty} \frac{\langle \nabla\phi(x(t)), \dot{x}(t)\rangle}{\|\nabla\phi(x(t))\| \cdot \|\dot{x}(t)\|} &= -\lim_{t\to+\infty} \frac{x_2^4 + x_3^2}{\sqrt{x_1^2 + x_2^2 + 1} \cdot \sqrt{x_2^6 + x_3^4}} \\
&= -\lim_{t\to+\infty} \frac{\frac{1}{4(t+k_1)^2} + \frac{1}{(t+k_2)^2}}{\sqrt{1 + \frac{1}{2(t+k_1)} + 1} \cdot \sqrt{\frac{1}{8(t+k_1)^3} + \frac{1}{(t+k_2)^4}}} \\
&= 0.
\end{aligned}$$

In this thesis, we get the convergence of ODE (3.4) by proving that there is an isolated point in the limit set $\Omega_3(x^0)$.

Theorem 3.14. *The limit set $\Omega_3(x^0)$ only contains a single point.*

Proof. Let $y(x)$, $z(x)$ be defined in (3.21), (3.22) respectively. From Theorem (3.10) and Theorem (3.11), $y(x)$ and $z(x)$ are two constants on $\Omega_3(x^0)$. For simplicity, these two constants are denoted by y^* and z^* respectively. For any $x \in \Omega_3(x^0)$, the following equalities hold

$$\begin{cases} Qx - A^T y^* = z^* - c, \\ Ax = b. \end{cases} \tag{3.32}$$

Since $\Omega_3(x^0)$ is nonempty, assume $x^* \in \Omega_3(x^0)$ and the number of its nonzero components is maximum for all $x \in \Omega_3(x^0)$. The index set $\{1, \ldots, n\}$ can be divided into two disjoint sets B, N based on the following rule

$$x_i^* > 0 \ (i \in B) \text{ and } x_i^* = 0 \ (i \in N).$$

If $B = \emptyset$, then there exists a unique point $x^* = 0$ in $\Omega_3(x^0)$. So suppose B is nonempty. Without loss of generality, furthermore suppose

$$B = \{1, \ldots, k\} \text{ and } N = \{k+1, \ldots, n\}.$$

Similarly, we can partition $z^* = \begin{pmatrix} z_B^* \\ z_N^* \end{pmatrix}$, $c = \begin{pmatrix} c_B \\ c_N \end{pmatrix}$, $e = \begin{pmatrix} e_B \\ e_N \end{pmatrix}$ respectively, where z_B^*, c_B, $e_B \in R^k$, and z_N^*, c_N, $e_N \in R^{(n-k)}$. Correspondingly, Q and A can be rewritten as

$$Q = \begin{pmatrix} Q_{BB} & Q_{BN} \\ Q_{BN}^T & Q_{NN} \end{pmatrix} \qquad A = \begin{pmatrix} A_B & A_N \end{pmatrix},$$

where $Q_{BB} \in R^{k\times k}$, $Q_{BN} \in R^{k\times(n-k)}$, $Q_{NN} \in R^{(n-k)\times(n-k)}$, $A_B \in R^{m\times k}$, and $A_N \in R^{m\times(n-k)}$.

Since x^* satisfies the equations (3.32), we have

$$\begin{cases} Q_{BB}x_B^* + Q_{BN}x_N^* - A_B^T y^* = z_B^* - c_B, \\ Q_{BN}^T x_B^* + Q_{NN}x_N^* - A_N^T y^* = z_N^* - c_N, \\ -A_B x_B^* - A_N x_N^* = -b. \end{cases} \tag{3.33}$$

It follows readily from $X^* z^* = 0$, $x_B^* > 0$ that $z_B^* = 0$, as a result

$$\begin{cases} Q_{BB}x_B^* - A_B^T y^* = -c_B, \\ Q_{BN}^T x_B^* - A_N^T y^* = z_N^* - c_N, \\ -A_B x_B^* = -b, \end{cases} \tag{3.34}$$

thus $c_B \in \text{range}\begin{pmatrix} Q_{BB} & -A_B^T \end{pmatrix}$.

For notational and analytical simplicity, vector $g \in R^{(m+n)}$ and matrix $W \in R^{(m+n)\times k}$ are introduced

$$g = \begin{pmatrix} -c_B + A_B^T y^* \\ A_N^T y^* + z_N^* - c_N \\ -b \end{pmatrix}, \quad W = (w_{ij})_{(m+n)\times k} = \begin{pmatrix} w_1 & w_2 & \dots & w_k \end{pmatrix} = \begin{pmatrix} Q_{BB} \\ Q_{BN}^T \\ -A_B \end{pmatrix},$$

where $w_i \in R^{(m+n)}$ $(i = 1, \dots, k)$.

Let $\delta_1 = \frac{1}{2}\min_{i\in B}\{x_i^*\}$. From the definition of x^*, we have

$$x_B > 0,\ x_N = 0,\ \forall x \in \Omega_3(x^0) \cap U(x^*, \delta_1), \tag{3.35}$$

where $U(x^*, \delta_1)$ is the δ_1-neighborhood of x^*. Next we will prove that x^* is an isolated point of $\Omega_3(x^0)$.

Combing (3.32) and (3.35), for any point $x = \begin{pmatrix} x_B \\ x_N \end{pmatrix} \in \Omega_3(x^0) \cap U(x^*, \delta_1)$, the following equality holds

$$W x_B = g, \tag{3.36}$$

(3.36) can be rewritten in an equivalent way as

$$\begin{cases} w_{11}x_1 + \cdots + w_{1k}x_k = g_1 \\ w_{21}x_1 + \cdots + w_{2k}x_k = g_2 \\ \vdots \\ w_{(m+n)1}x_1 + \cdots + w_{(m+n)k}x_k = g_{(m+n)} \end{cases} \tag{3.37}$$

or

$$x_1 w_1 + \cdots + x_k w_k = g. \tag{3.38}$$

If $\text{rank}(w_1, w_2, \ldots, w_k) = k$, g will be expressed uniquely as a linear combination of $w_1, w_2, \ldots, w_k$. Thus except for x^*, there is no other point in $\Omega_3(x^0) \cap U(x^*, \delta_1)$ such that (3.38) holds. In other words, x^* is an isolated point of $\Omega_3(x^0)$.

If $\text{rank}(w_1, w_2, \ldots, w_k) = r < k$, we notice that

$$\begin{aligned} \frac{dx}{dt} &= -X[I_n - XA^T(AX^2A^T)^{-1}AX]X(Qx + c) \\ &= -X^2\Big[\begin{pmatrix} Q & -A^T \end{pmatrix}\begin{pmatrix} x \\ y(x) \end{pmatrix} + c\Big] \\ &= -\begin{pmatrix} X_B^2 & 0 \\ 0 & X_N^2 \end{pmatrix}\left[\begin{pmatrix} Q_{BB} & Q_{BN} & -A_B^T \\ Q_{BN}^T & Q_{NN} & -A_N^T \end{pmatrix}\begin{pmatrix} x_B \\ x_N \\ y(x) \end{pmatrix} + \begin{pmatrix} c_B \\ c_N \end{pmatrix}\right], \end{aligned}$$

thus

$$\frac{dx_B}{dt} = -X_B^2\Big[\Big(Q_{BB} \quad Q_{BN} \quad -A_B^T\Big)\begin{pmatrix} x_B \\ x_N \\ y(x) \end{pmatrix} + c_B\Big] \tag{3.39}$$

$$= -X_B^2\Big[W^T\begin{pmatrix} x \\ y(x) \end{pmatrix} + c_B\Big]. \tag{3.40}$$

Based on the definition of x^* and the boundedness of the optimal solution set, the case that $r = 0$ is excluded. If $r = 0$, i.e., $w_1 = w_2 = \cdots = w_k = 0$, $\frac{dx_i(t)}{dt}$ $(i = 1, \ldots, k)$ will reduce to

$$\frac{dx_i(t)}{dt} = -c_i x_i^2.$$

By Theorem 3.6, $\lim_{t\to+\infty} \frac{dx_i(t)}{dt} = 0$. For any $i \in \{1, \ldots, k\}$, there are two cases

(a) $c_i = 0 \Rightarrow x_i(t) \equiv x_i^0$, but $W = 0$ and $c_i = 0$ imply the optimal solution set of problem (P_4) is unbounded, which contradicts with Assumption 3.1;

(b) $c_i \neq 0 \Rightarrow x_i^* = 0$, it is a contradiction to the assumption that x_i^* is positive.

Thus $1 \le r < k$, without loss of generality, assume $\{w_1, w_2, \ldots, w_r\}$ is a maximum linearly independent group of $\{w_1, w_2, \ldots, w_k\}$. Thus there exists a matrix $V = (v_{ij})_{r\times(k-r)} \in R^{r\times(k-r)}$ such that

$$w_{r+i} = \sum_{j=1}^{r} v_{ji} w_j, \; i = 1, \ldots, k-r. \tag{3.41}$$

Denote

$$U = \Big(u_1 \quad u_2 \quad \ldots \quad u_{k-r}\Big) = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1(k-r)} \\ v_{21} & v_{22} & \cdots & v_{2(k-r)} \\ \vdots & \vdots & \cdots & \vdots \\ v_{r1} & v_{r2} & \cdots & v_{r(k-r)} \\ -1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & -1 \end{pmatrix},$$

where $u_i \in R^k$ $(i = 1, 2, \ldots, k-r)$, thus

$$WU = \begin{pmatrix} w_1 & w_2 & \ldots & w_k \end{pmatrix} \begin{pmatrix} u_1 & u_2 & \ldots & u_{k-r} \end{pmatrix} = 0. \tag{3.42}$$

Since $c_B \in \text{range}\begin{pmatrix} Q_{BB} & -A_B^T \end{pmatrix} \subseteq \text{range}(W^T)$, we get

$$u_i^T c_B = 0, \ i = 1, \ldots, k-r. \tag{3.43}$$

To overcome the difficulty caused by the degeneracy in linear equations (3.37), let us define

$$y_i(t) = u_i^T X_B^{-1}(t) e_B, \ i = 1, \ldots, k-r, \ t \geq 0, \tag{3.44}$$

where $x(t)$ is the solution of ODE (3.4). Theorem 3.3 and Theorem 3.4 indicate (3.44) is well defined on $[0, \ +\infty)$. From (3.42) and (3.43), we have

$$\begin{aligned} \frac{dy_i(t)}{dt} &= u_i^T X_B^{-2} X_B^2 \left[\begin{pmatrix} Q_{BB} & Q_{BN} & -A_B^T \end{pmatrix} \begin{pmatrix} x_B \\ x_N \\ y(x) \end{pmatrix} + c_B \right] \\ &= u_i^T W^T \begin{pmatrix} x \\ y(x) \end{pmatrix} + u_i^T c_B \\ &= (W u_i)^T \begin{pmatrix} x \\ y(x) \end{pmatrix} + u_i^T c_B \\ &\equiv 0, \quad i = 1, \ldots, k-r. \end{aligned}$$

Thus there exist $k-r$ constants $\bar{c}_i$ $(i = 1, \ldots, k-r)$ such that

$$y_i(t) \equiv \bar{c}_i, \ i = 1, \ldots, k-r, \ t \in [0, \ +\infty). \tag{3.45}$$

Clearly, the following nonlinear equations hold

$$u_i^T X_B^{-1} e_B \equiv \bar{c}_i, \ i = 1, \ldots, k-r, \ \forall x \in \Omega_3(x^0) \cap U(x^*, \delta_1). \tag{3.46}$$

Let $H \in R^{r\times k}$ be a matrix generated by choosing r linearly independent rows, say $l_1, \ldots, l_r$, from matrix W. H can be denoted as

$$H = \begin{pmatrix} w_{l_1 1} & \cdots & w_{l_1 r} & w_{l_1 (r+1)} & \cdots & w_{l_1 k} \\ w_{l_2 1} & \cdots & w_{l_2 r} & w_{l_2 (r+1)} & \cdots & w_{l_2 k} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ w_{l_r 1} & \cdots & w_{l_r r} & w_{l_r (r+1)} & \cdots & w_{l_r k} \end{pmatrix}$$

We consider the following nonlinear equations

$$\begin{cases} w_{l_1 1} s_1 + \cdots + w_{l_1 r} s_r + w_{l_1 (r+1)} s_{r+1} + w_{l_1 (r+2)} s_{r+2} + \cdots + w_{l_1 k} s_k = g_{l1} \\ w_{l_2 1} s_1 + \cdots + w_{l_2 r} s_r + w_{l_2 (r+1)} s_{r+1} + w_{l_2 (r+2)} s_{r+2} + \cdots + w_{l_2 k} s_k = g_{l2} \\ \vdots \\ w_{l_r 1} s_1 + \cdots + w_{l_r r} s_r + w_{l_r (r+1)} s_{r+1} + w_{l_r (r+2)} s_{r+2} + \cdots + w_{l_r k} s_k = g_{lr} \\ \frac{v_{11}}{s_1} + \cdots + \frac{v_{r1}}{s_r} - \frac{1}{s_{r+1}} - 0 - \cdots - 0 = \bar{c}_1 \\ \frac{v_{12}}{s_1} + \cdots + \frac{v_{r2}}{s_r} - 0 - \frac{1}{s_{r+2}} - \cdots - 0 = \bar{c}_2 \\ \vdots \\ \frac{v_{1(k-r)}}{s_1} + \cdots + \frac{v_{r(k-r)}}{s_r} - 0 - 0 - \cdots - \frac{1}{s_k} = \bar{c}_{k-r} \end{cases} \tag{3.47}$$

From (3.37) and (3.46), for any $x \in \Omega_3(x^0) \cap U(x^*, \delta_1)$, $s = x_B = (x_1, x_2, \ldots, x_k)^T$ is a solution of the system (3.47). The Jacobian matrix of the nonlinear equations (3.47) is

$$J(s) = \begin{pmatrix} w_{l_1 1} & \cdots & w_{l_1 r} & w_{l_1 (r+1)} & w_{l_1 (r+2)} & \cdots & w_{l_1 k} \\ w_{l_2 1} & \cdots & w_{l_2 r} & w_{l_2 (r+1)} & w_{l_2 (r+2)} & \cdots & w_{l_2 k} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{l_r 1} & \cdots & w_{l_r r} & w_{l_r (r+1)} & w_{l_r (r+2)} & \cdots & w_{l_r k} \\ \frac{-v_{11}}{s_1^2} & \cdots & \frac{-v_{r1}}{s_r^2} & \frac{1}{s_{r+1}^2} & 0 & \cdots & 0 \\ \frac{-v_{12}}{s_1^2} & \cdots & \frac{-v_{r2}}{s_r^2} & 0 & \frac{1}{s_{r+2}^2} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{-v_{1(k-r)}}{s_1^2} & \cdots & \frac{-v_{r(k-r)}}{s_r^2} & 0 & 0 & \cdots & \frac{1}{s_k^2} \end{pmatrix}.$$

From (3.41), after a series of Gaussian operations, $J(s)$ can be converted into

$$\begin{aligned}
\bar{J}(s) &= \begin{pmatrix}
w_{l_1 1} & \cdots & w_{l_1 r} & 0 & 0 & \cdots & 0 \\
w_{l_2 1} & \cdots & w_{l_2 r} & 0 & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
w_{l_r 1} & \cdots & w_{l_r r} & 0 & 0 & \cdots & 0 \\
\frac{-v_{11}}{s_1^2} & \cdots & \frac{-v_{r1}}{s_r^2} & \frac{1}{s_{r+1}^2} + \sum_{j=1}^{r} \frac{v_{j1}^2}{s_j^2} & \sum_{j=1}^{r} \frac{v_{j2}v_{j1}}{s_j^2} & \cdots & \sum_{j=1}^{r} \frac{v_{j(k-r)}v_{j1}}{s_j^2} \\
\frac{-v_{12}}{s_1^2} & \cdots & \frac{-v_{r2}}{s_r^2} & \sum_{j=1}^{r} \frac{v_{j1}v_{j2}}{s_j^2} & \frac{1}{z_{r+2}^2} + \sum_{j=1}^{r} \frac{v_{j2}^2}{s_j^2} & \cdots & \sum_{j=1}^{r} \frac{v_{j(k-r)}v_{j2}}{s_j^2} \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{-v_{1(k-r)}}{s_1^2} & \cdots & \frac{-v_{r(k-r)}}{s_r^2} & \sum_{j=1}^{r} \frac{v_{j1}v_{j(k-r)}}{s_j^2} & \sum_{j=1}^{r} \frac{v_{j2}v_{j(k-r)}}{s_j^2} & \cdots & \frac{1}{s_k^2} + \sum_{j=1}^{r} \frac{v_{j(k-r)}^2}{s_j^2}
\end{pmatrix} \\
&\doteq \begin{pmatrix} M_1 & 0 \\ M_2 & M_3 \end{pmatrix},
\end{aligned}$$

where $M_1 \in R^{r\times r}$, $M_2 \in R^{(k-r)\times r}$, and $M_3 \in R^{(k-r)\times(k-r)}$. It is very easy to verify that

(a) $rank(M_1) = r$, and

(b)

$$\begin{aligned}
M_3 &= \begin{pmatrix}
\frac{v_{11}}{s_1} & \frac{v_{21}}{s_2} & \cdots & \frac{v_{r1}}{s_r} \\
\frac{v_{12}}{s_1} & \frac{v_{22}}{s_2} & \cdots & \frac{v_{r2}}{s_r} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{v_{1(k-r)}}{s_1} & \frac{v_{2(k-r)}}{s_2} & \cdots & \frac{v_{r(k-r)}}{s_r}
\end{pmatrix}
\begin{pmatrix}
\frac{v_{11}}{s_1} & \frac{v_{12}}{s_1} & \cdots & \frac{v_{1(k-r)}}{s_1} \\
\frac{v_{21}}{s_2} & \frac{v_{22}}{s_2} & \cdots & \frac{v_{2(k-r)}}{s_2} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{v_{r1}}{s_r} & \frac{v_{r2}}{s_r} & \cdots & \frac{v_{r(k-r)}}{s_r}
\end{pmatrix} \\
&+ \begin{pmatrix}
\frac{1}{s_{r+1}^2} & 0 & \cdots & 0 \\
0 & \frac{1}{s_{r+2}^2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{1}{s_k^2}
\end{pmatrix}.
\end{aligned}$$

Thus $J(s)$ is invertible $\forall s\ (\in R^k)\ > 0$.

Let $F(s) = 0$ be the system (3.47). From previous discussions, we know (i) $\forall x \in \Omega_3(x^0) \cap U(x^*, \delta_2)$, $s = (x_1, x_2, \ldots, x_k)^T$ is a solution of $F(s) = 0$, in particular,

$s^* = (x_1, x_2, \ldots, x_k)^T$ is also a solution of $F(s) = 0$; (ii) $\frac{\partial F}{\partial s}$ is invertible $\forall s\ (\in R^k)\ > 0$; and (iii) $s^* > 0$. By Lemma 1.6, $s = s^*$ must be an isolated point satisfying $F(s) = 0$. Therefore, there exists a $\delta_2 > 0$ ($\delta_2 \leq \delta_1$) such that for any $x \in \Omega_3(x^0) \cap U(x^*, \delta_2)$, $s = (x_1, x_2, \ldots, x_k)^T$ is a solution of system (3.47) if and only if $x = x^*$. Thus there is only one point $x^* \in \Omega_3(x^0) \cap U(x^*, \delta_2)$, i.e., x^* is an isolated point of $\Omega_3(x^0)$. But $\Omega_3(x^0)$ is connected, thus there is only one point x^* in $\Omega_3(x^0)$. The proof is complete. □

Theorem 3.13 and Theorem 3.14 ensure the solution trajectory of the system (3.4) will tend to an optimal solution of the problem (P_4) as $t \to +\infty$. Thus the strong convergence is proved.

3.6 Concluding Remarks

In this chapter, by adopting the continuous method framework, the first-order affine scaling continuous trajectory is discussed. The convergence of the continuous trajectory is obtained. In fact, the properties outlined in this chapter also hold for the following ODE system, such as existence, uniqueness and convergence of the solution trajectory, optimality of the limit point.

$$\frac{dx}{dt} = -X^\gamma P_{AX^\gamma} X^\gamma (Qx + c), \quad x(0) = x^0 \in \mathcal{P}^{++}. \tag{3.48}$$

where $\gamma \geq \frac{1}{2}$ is constant. Lyapunov direct method is a powerful tool to prove the convergence of ODE system. But for complicated system (3.4), it is very difficult to construct a proper Lyapunov function to prove the convergence. It is an open problem whether the ODE system (3.4) is stable in the sense of Lyapunov.

Chapter 4

Summary

4.1 Summary of the Thesis

In this thesis, we propose several continuous method models to solve CQP problems with different types of constraints. Starting from any interior point, the solution trajectories of our continuous method models are always restricted in the interior of the feasible region. Moreover, the proofs of convergence are different from those appeared in neural network, where Lyapunov potential functions are needed.

For the nonnegativity and box constrained CQP, without projection technology, the structures of our ODE systems are simpler and sufficiently smooth. In the future, it is possible that more effective numerical solution schemes may be explored for our ODE systems. Large-scale random problems illustrate that our new methods are very encouraging.

For CQP in the standard form, the first-order affine scaling continuous trajectory is considered. Two existing important theoretical tools for strong convergence in linear programming are not suitable for quadratic programming, a new method to prove the strong convergence is presented here. In fact, the proof skills in this thesis can be generalized to prove the convergence of the following dynamical system

$$\frac{dx}{dt} = -[I_n - A^T(AA^T)^{-1}A](Qx + c), \tag{4.1}$$

which can be used to solve the following equality constrained CQP

$$\begin{array}{ll} \min & q(x) = \frac{1}{2}x^TQx + c^Tx \\ s.t. & Ax = b. \end{array} \tag{P_7}$$

4.2 Future Research

Compared with the well-known quadratic programming, the applications of fractional programming are less known. The fractional programming problems are particularly useful in the solution of economic problems in which various activities use certain resources in various pro-portions, while the objective is to optimize a certain indicator, usually the most favorable return-on-allocation ratio subject to the constraint imposed on the availability of goods. Especially, linear fractional programming arises in fields of game theory, network flows; quadratic fractional programming is used on field production planning and inventories. For more applications, see [9, 18, 86]. As it is known, the ratio of convex and concave functions is not convex in general, even numerator and denominator are both linear. It is possible to solve quasi-convex fractional programming by several of the standard convex programming algorithms. In future study, we plan to present new continuous method models for solving fractional programming.

Effective numerical solution schemes for our ODEs are urgently needed. For optimization problems, since we are only interested in the limit point of the ODE, it is possible to introduce some new numerical method framework so that we can find an optimal solution in less time. Taking the ODE system (2.9) for example, we present our idea to find an optimal solution briefly. ODE system (2.9) can be rewritten as

$$\begin{cases} \frac{dx(t)}{dt} = -X(Qx+c), \ t \geq 0, \\ x(0) = x^0 > 0. \end{cases} \tag{4.2}$$

For ODE (4.2), we introduce the following implicit-explicit Euler iterative scheme

$$x_{k+1} = x_k - h_k X_{k+1}(Qx_k + c), \tag{4.3}$$

where x_k is the k-th iterative point, and h_k is the corresponding step size to be determined. For simplicity, scheme (4.3) can be rewritten in the following equivalent form

$$[x_{k+1}]_i = \frac{[x_k]_i}{1 + h_k[Qx_k + c]_i}, \;\; i = 1, 2, \ldots, n. \tag{4.4}$$

Considering the original quadratic programming problem, we require $x_k > 0$ and $q(x_{k+1}) \le q(x_k)$ for any k.

Claim 4.1. *For iterative scheme (4.4), $x_{k+1} > 0$ if step size h_k is selected by the following rule*

$$h_k = \begin{cases} 2h_{k-1}, & \text{if } \min\{[Qx_k + c]_i\} > 0; \\ r, & \text{if } -1 < \min\{[Qx_k + c]_i\} < 0; \\ \frac{-0.99}{\min\{[Qx_k+c]_i\}}, & \text{if } \min\{[Qx_k + c]_i\} \le -1; \end{cases} \tag{4.5}$$

where $0 < r < 1$. In our following numerical simulation, we take $r = 0.8$.

From (4.3), we have

$$\begin{aligned} q(x_{k+1}) &= \frac{1}{2}x_{k+1}^T Qx_{k+1} + c^T x_{k+1} \\ &= \frac{1}{2}[x_k - h_k X_{k+1}(Qx_k + c)]^T Q[x_k - h_k X_{k+1}(Qx_k + c)] \\ &\quad + c^T[x_k - h_k X_{k+1}(Qx_k + c)] \\ &= \frac{1}{2}x_k^T Qx_k + c^T x_k - h_k(Qx_k + c)^T X_{k+1}(Qx_k + c) \\ &\quad + \frac{h_k^2}{2}(Qx_k + c)^T X_{k+1} Q X_{k+1}(Qx_k + c) \\ &= q(x_k) - \frac{h_k}{2}(Qx_k + c)^T X_{k+1}(2X_{k+1}^{-1} - h_k Q)X_{k+1}(Qx_k + c). \end{aligned}$$

If $(2X_{k+1}^{-1} - h_k Q)$ is positive definite, then $q(x_{k+1}) \le q(x_k)$. To ensure this result, we only need

$$\min\{\frac{2(1 + h_k[Qx_k + c]_i}{h_k[x_k]_i}\} \ge \min\{\frac{2}{h_k[x_k]_i}\} + \min\{\frac{2[Qx_k + c]_i}{[x_k]_i}\} > \lambda_{\max}(Q), \tag{4.6}$$

where $\lambda_{\max}(Q)$ is the maximum eigenvalue of Q. Thus the following claim is clear.

Claim 4.2. *For iterative scheme (4.4), $q_{k+1} \le q_k$ if step size h_k is selected by the following rule*

$$h_k = \begin{cases} \textit{any positive number}, & \text{if } \lambda_{\max}(Q) - \gamma \le 0; \\ \frac{2}{(\lambda_{\max}(Q)-\gamma)\max\{[x_k]_i\}}, & \text{if } \lambda_{\max}(Q) - \gamma > 0. \end{cases} \tag{4.7}$$

where $\gamma = 2\min\{[X_k^{-1}(Qx_k + c)]_i\}$.

Let $\tilde{h}_k$ and $\bar{h}_k$ be determined by rules (4.5) and (4.7) respectively, furthermore $h_k = \min\{\tilde{h}_k, \bar{h}_k\}$. Then h_k will be an appropriate step size for iterative scheme (4.4). To illustrate the feasibility of iterative scheme (4.4), ten random nonnegativity constrained CQP problems (generated as in Section 2.4 in Chapter 2) are tested. For consistency, all the parameters are the same to those in Section 2.4.1. The initial points for the ten randomly generated (Q, c)s are also set to $x^0 = (1, \cdots, 1)^T$. For stopping criterion $|X_k(Qx_k + c)|_\infty \leq 10^{-4}$, the numerical results are reported in Table 4.1. For better understanding of iterative scheme (4.4), Figure 4.1 depicts the transient behavior of step size h_k with $n = 1000$ and $\epsilon = 10^{-5}$.

Table 4.1: The performance of iterative scheme (4.4) ($\epsilon = 10^{-4}$)

n	200	400	600	800	1000	1200	1400	1600	1800	2000
CPU(s)	0.04	0.03	0.05	0.04	0.06	0.08	0.10	0.12	0.16	0.19

From Table 2.1 and Table 4.1, we can see that the iterative scheme (4.4) is faster in finding an optimal solution for problem (P_2) in moderate accuracy. It is natural since the iterative points are not required to be close to the real solution trajectory. It should be noted that the iterative scheme (4.4) is very sensitive to the stopping criterion. Figure 4.1 indicts that step size h_k will tend to a constant. The investigation on various theoretical results for iterative scheme (4.4) is one of our future goals.

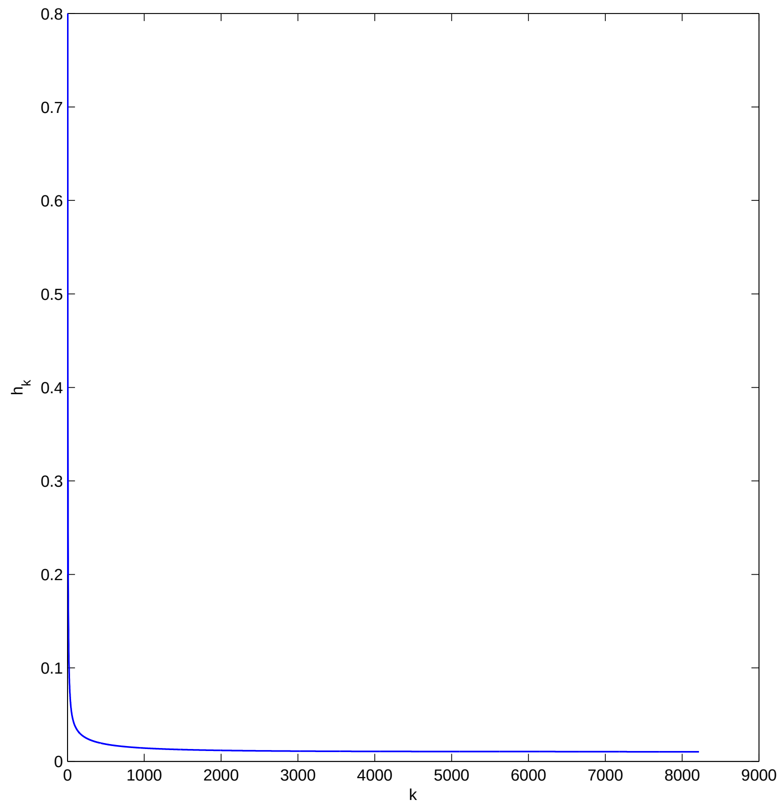

Figure 4.1: The transient behavior of step size h_k for the iterative scheme (4.4) with $n = 1000$ and $\epsilon = 10^{-5}$.

Bibliography

[1] P.-A. Absil, R. Mahony and B. Andrews, *Convergence of the iterates of descent methods for analytic cost functions*, SIAM Journal on Optimization, 16 (2005), pp. 531-547.

[2] P.-A. Absil and K. Kurdyka, *On the stable equilibrium points of gradient sys- tems*, Systems & Control Letters, 55 (2006), pp. 573-577.

[3] I. Adler, N. Karmarkar, M. G. C. Resend and G. Veiga, *An implementation of Karmarkars algorithm for linear programming*, Mathematical Programming, 44 (1989), pp. 297-335.

[4] I. Adler and R. D. C. Monteiro, *Limiting behavior of the affine scaling continuous trajectories for linear programming problems*, Mathematical Programming, 50 (1991), pp. 29-51.

[5] A. S. Antipin, *Minimization of convex functions on convex sets by means of differential equations*, Differential Equations, 30 (1994), pp. 1365-1375.

[6] D. A. Bayer and J. C. Lagarias, *The nonlinear geometry of linear programming. I: Affine and projective scaling trajectories*, Transactions of the American Mathematical Society, 314 (1989), pp. 499-526.

[7] D. A. Bayer and J. C. Lagarias, *The nonlinear geometry of linear programming. II: Legendre transform coordinates and central trajectories*, Transactions of the American Mathematical Society, 314 (1989), pp. 527-581.

[8] M. S. Bazaraa, H. D. Sherali and C. M. Shetty, *Nonlinear programming, theory and algorithms*, John Wiley & Sons, New York, NY, 1993.

[9] G. R. Bitrang and A. G. Novaesa, *Linear programming with a fractional objective function*, Operations Research, 21 (1973), pp. 22-29.

[10] A. A. Brown and M.C. Bartholomew-Biggs, *Some effective methods for uncon- strained optimization based on the solution of systems of ordinary differential equations*, Journal of Optimization Theory and Applications, 62 (1988), pp. 211-224.

[11] C. A. Botsaris, *Differential gradient methods*, Journal of Mathematical Analysis and Applications, 63 (1978), pp. 177-198.

[12] M. T. Chu, *On the continuous realization of the iterative processes*, SIAM Review, 30 (1988), pp. 375-387.

[13] M. T. Chu and M. M. Lin, *Dynamical system characterization of the central path and its variants-a revisit*, SIAM Journal on Applied Dynamical Systems, 10 (2011), pp. 887-905.

[14] A. Cichocki and R. Unbehauen, *Neural networks for optimization and signal processing*, London, United Kingdom, Wiley, 1993.

[15] E. A. Coddington and N. Levinson, *Theory of ordinary differential equations*, McGraw-Hill Book Co., New York, 1955.

[16] C. Y. Dang and L. Xu, *A barrier function method for the nonconvex quadratic programming problem with box constraints*, Journal of Global Optimization, 18 (2000), pp. 165-188.

[17] R. S. Dembo and T. Ulrich, *On the minimization of quadratic functions subject to box constrains*, 1984, http://www.cs.yale.edu/publications/techreports/tr302.pdf.

[18] C. Derman, *On sequential decisions and markov chains*, Management Science, 9 (1962), pp. 16-24.

[19] I. DIENER, *On the global convergence of path-following methods to determine all solutions to a system of nonlinear equations*, Mathematical Programming, 39 (1987), pp. 181-188.

[20] I. I. DIKIN, *Iterative solution of problems of linear and quadratic programming*, Soviet Mathematics Doklady, 8 (1967), pp. 674-675 (in Russian).

[21] I. I. DIKIN, *On the convergence of an iterative process*, Upravlyaemye Sistemy, 12 (1974), pp. 54-60 (in Russian).

[22] W. DINKELBACH, *On nonlinear fractional programming*, Management Science, Series A, 13 (1967), pp. 492-498.

[23] L. M. G. DRUMMOND AND B. F. SVAITER, *On well definedness of the central path*, Journal of Optimization Theory and Applications, 102 (1999), pp. 223-237.

[24] B. C. EAVES, *On the basic theorem of complementarity*, Mathematical Programming, 1 (1971), pp. 68-75.

[25] A. V. FIACCO AND G. P. MCCORMICK, *Nonlinear programming: Sequential unconstrained minimization techniques*, SIAM, Philadelphia, USA, 1990.

[26] R. FLETCHER, *Practical methods of optimization* (2nd edtition), John Wiley and Sons, Chichester, New York, Brisbance, Toronto and Singapore, 1987.

[27] A. FRIEDLANDER, J. M. MARTINEZ AND M. RAYDAN, *A new method for large-scale box con strained convex quadratic minimization problems*, Optimization Methods and Software, 5 (1995), pp. 57-74.

[28] X. B. GAO, L.-Z. LIAO AND W. M. XUE, *A neural network for a class of convex quadratic minimax problems with constraints*, IEEE Transactions on Neural Networks, 16 (2004), pp. 622-628.

[29] D. M. Gay, *A variant of Karmarkars linear programming algorithm for problem sin standard form*, Mathematical Programming, 37 (1987), pp. 81-90.

[30] A. M. Geoffrion, *Strictly concave parametric programming, part I: Basic theory*, Management Science, 13 (1967), pp. 244-253.

[31] A. M. Geoffrion, *Strictly concave parametric programming, part II: Additional theory and computational considerations*, Management Science, 13 (1967), pp. 359-370.

[32] P. E. Gill, W. Murrray and M. H. Wright, *Practical optimization*, Academic Press, London and New York, 1981.

[33] C. C. Gonzaga and L. A. Carlos, *A primal affine-scaling algorithm for linearly constrained convex programs*, 2002, http://www.optimization-online.org/DB_HTML/2002/09/531.html.

[34] Q. M. Han, L.-Z. Liao, H. D. Qi and L. Q. Qi, *Stability analysis of gradient- based neural networks for optimization problems*, Journal of Global Optimization, 19 (2001), pp. 363-381.

[35] D. den Hertog and C. Roos, *A survey of search directions in interior point methods for linear programming*, Mathematical Programming, 52 (1991), pp. 481-509.

[36] A. Jeflea, *A parametric study for solving nonlinear fractional problems*, An St. Univ Ovidius constanta, 11 (2003), pp. 87-92.

[37] M. C. Joshi, *Ordinary differential equations: modern perspective*, Alpha Science International Ltd., 2006.

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

[39] M. Kojima, N. Megiddo and Y. Ye, *An interior point potential reduction algorithm for the linear complementarity problem*, Mathematical Programming, 54 (1992), pp. 267-279.

[40] C. L. Lawson and R. J. Hanson. *Solving least squares problems*, Englewood Cliffs, NJ: Prentice-hall, 1974.

[41] L.-Z. Liao, *A continuous method for convex programming problems*, Journal of Opimization Theory and Applications, 124 (2005), pp. 207-226.

[42] L.-Z. Liao, H. D. Qi and L. Q. Qi, *Neurodynamical optimization*, Journal of Global Optimization, 28 (2004), pp. 175-195.

[43] V. Losert and E. Akin, *Dynamics of games and genes: disrete versus contibuous time*, Journal of Mathematical Biology, 17 (1983), pp. 241-251.

[44] O. L. Mangasarian, *A simple characterization of solution sets of convex programs*, Operations Research Letters, 7 (1988), pp. 21-26.

[45] N. Megiddo and M. Shub, *Boundary behavior of interior point algorihms for linear programming*, Mathematics of Operations Research, 14 (1989), pp. 97-146.

[46] R. D. C. Monteiro and I. Adler, *Interior path following primal-dual algorithms. part I: Linear programming*, Mathematical Programming, 44 (1989), pp. 27-41.

[47] R. D. C. Monteiro and I. Adler, *Interior path following primal-dual algorithms. part II: Convex quadratic programming*, Mathematical Programming, 44 (1989), pp. 43-66.

[48] R. D. C. Monteiro, I. Adler and M. G. C. Resende, *A polynomial-time primal-dual affine scaling algorithm for linear and convex quadratic program-*

ming and its power series extension, Mathematics of Operations Research, 15 (1990), pp. 191-214.

[49] R. D. C. Monteiro, *On the continuous trajectories for a potential reduction algorithm for linear programming*, Mathematics of Operations Research, 17 (1992), pp. 225-253.

[50] R. D. C. Monteiro, *Convergence and bounary behavior of the projective scaling trajectories for linear programmin*, Mathematics of Operations Research, 16 (1991), pp. 842-858.

[51] R. D. C. Monteiro and T. Tsuchiya, *Global convergence of the affine scaling algorothm for convex quadratic programming*, SIAM Journal on Optimization, 8 (1998), pp. 26-58.

[52] R. D. C. Monteiro and F. J. Zhou, *On the existence and convergence of the central path for convex programming and some duality results*, Computational Optimization and Applications, 10 (1998), pp. 51-77.

[53] J. J. Moŕe and T. Gerardo, *Algorithms for bound constrained quadratic programming problems*, Numerische Mathematik, 55 (1989), pp. 377-400.

[54] J. J. Moŕe and G. Toraldo, *On the solution of large quadratic programming problems with bound constrains*, SIAM Journal on Optimization, 1 (1991), pp. 93-113.

[55] P. M. Pardalos and S. A. Vavasis, *Quadratic programming with one negative eigenvalue is NP-hard*, Journal of Global Optimization, 1 (1991), pp. 15-22.

[56] K. Ritter, *Ein Verfahren zur Lösung parameterabhängiger, nichtlinearer maximum-probleme*, Unternehmensforschung, 6 (1962), pp. 149-166.

[57] R. T. Rockafellar, *Convex analysis*, Princeton University Press, Princeton, New Jersey, 1970.

[58] W. Rudin, *Principles of mathematical analysis*, third ed., McGraw-Hill Book Company, 1976.

[59] R. Saigal, *A simple proof of a primal affine scaling method*, Annals of Operations Research, 62 (1996), pp. 303-324.

[60] F. Sha, L. K. Saul and D. D. Lee, *Multiplicative updates for nonnegative quadratic programming in support vector machines*, In Advances in Neural Information Processing Systems, 15 (2003), pp. 1041-1048.

[61] J. J. E. Slotine and W. Li, *Applied nonlinear control*, Prentice Hall, New Jersey, 1991.

[62] S. Mehrotra and J. Sun, *A Method of Analytic Centers for Quadratically Constrained Convex Quadratic Programs*, SIAM Journal on Numerical Analysis, 28(1991), Issue 2, pp. 529-54.

[63] J. Sun, *A convergence proof for an affine scaling algorithm for convex quadratic programming without nondegeneracy assumptions*, Mathematical Programming, 60 (1993), pp. 69-79.

[64] J. Sun, *A convergence analysis for a convex version of Dikin's algorithm*, Annals of Operations Research, 62(1996), pp. 357-374.

[65] Y. Tan and C. Deng, *Solving for a quadratic programming with a quadratic constraint based on a neural network frame*, Neurocomputing, 30 (2000), pp. 117-127.

[66] P. Tseng and Z.-Q. Luo, *On the convergence of the affine-scaling algorithm*, Mathematical Programming, 56 (1992), pp. 301-319.

[67] P. Tseng, I. M. Bomze and W. Schachinger, *A first-order interior point method for linearly constrained smooth optimization*, Mathematical Programming, 127 (2011), pp. 399-424.

[68] T. Tsuchiya, *Global convergence property of the affine scaling methods for primal degenerate linear programming problems*, Mathematics of Operations Research, 17 (1992), pp. 527-557.

[69] T. Tsuchiya, *Affine scaling algorithm*, in: T. Terlaky, ed., Interior Point Methods of Mathematical Programming, Kluwer Academic Publishers, 1996, 35-82.

[70] I. I. Vrabie, *Diffenential equations: An introduction to basic concepts, results and applications*, World Scientific Publishing Co. Pte. Ltd, 2004.

[71] J. Wang, *Recurrent neural network for solving quadratic programming solving quadratic programming problems with equality constraints*, Electronics Letters, 28 (1992), pp. 1345-1347.

[72] J. Wang, *Analysis and design of a recurrent neural network for linear pro gramming*, IEEE Transactions on Circuits and Systems, 40 (1993), pp. 613-618.

[73] X. Y. Wu, Y. S. Xia, J. Li and W. K. Chen, *A high performance neural network for solving linear and quadratic programming problems*, IEEE transactions on Neural Networks, 7 (1996), pp. 643-651.

[74] Y. S. Xia, *A new neural network for solving linear programming problems and its applications*, IEEE ransactions on Neural Networks, 7 (1996), pp. 525-529.

[75] Y. S. Xia, *A new neural network for solving linear and quadratic programming problems*, IEEE transactions on Neural Networks, 7 (1996), pp. 1544-1547.

[76] Y. S. Xia and J. Wang, *A general methodology for designing globally convergent optimization neural networks*, IEEE transactions on Neural Networks, 9 (1998), pp. 1331-1343.

[77] Y. S. Xia and J. Wang, *On the stability of globally projected dynamical systems*, Journal of Optimization Theory and Applications, 106 (2000), pp. 129-150.

[78] Y. S. Xia and J. Wang, *Global asymptotic and axponential atability of a dynamic neural system with asymmetric connection weights*, IEEE transactions on Automatic Control, 46 (2001), pp. 635-638.

[79] Y. S. Xia, G. Feng and J. Wang, *A recurrent neural network with exponential convergence for solving convex quadratic program and related linear piecewise equations*, IEEE transactions on Neural Networks, 17 (2004), pp. 1003-1015.

[80] Y. S. Xia and J. Wang, *A recurrent neural network for solving convex programs subject to linear constraints*, IEEE transactions on Neural Networks, 16 (2005), pp. 379-386.

[81] Y. Ye, *Interior-point algorithms for global optimization*, Annals of Operations Research, 25 (1990), pp. 59-73.

[82] Y. Ye and E. Tse, *An extension of Karmarkar's projective algorithm for convex quadratic programming*, Mathematical Programming, 44 (1989), pp. 157-180.

[83] D. Zhang and A. Nagurney, *On the stability of projected dynamical systems*,Journal of Optimization Theory and Applications, 85 (1995), pp. 97-124.

[84] Q. J. Zhang and X. Q. Lu, *A recurrent neural network for nonlinear fractional programming*, Mathematical Problems in Engineering, 2012, http://www.hindawi.com/journals/mpe/2012/807656/.

[85] S. Zhang and A. G. Constantinides, *Lagrange programming neural networks*, IEEE Transactions on Circuits and Systems, 39 (1992), pp. 441-452.

[86] W. T. Ziemba, C. Parkan and R. B. Hill, *Calculation of investment portfolios with risk free borrowing and lending*, Management Science, 21 (1974), pp. 209-222.

Curriculum Vitae

Academic qualifications of the thesis author, Mr. YUE Hongwei:

- Received the degree of Bachelor of Science (Mathematics) from Henan Normal University (China), July 2003.

- Received the degree of Master of Science (Operation Research and Control Theory) from Southeast University (China), April 2006.

February 2014