UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

DIFFERENTIAL-ALGEBRAIC EQUATIONS IN PRIMAL DUAL INTERIOR POINT OPTIMIZATION METHODS: A NEW APPROACH TO THE PARAMETERIZATION OF THE CENTRAL TRAJECTORY

A Dissertation

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

Doctor of Philosophy

By

SUAT KASAP

Norman, Oklahoma

2003

UMI Number: 3082923

UMI®

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

DIFFERENTIAL-ALGEBRAIC EQUATIONS

IN PRIMAL DUAL INTERIOR POINT OPTIMIZATION METHODS:

A NEW APPROACH TO THE PARAMETRIZATION

OF THE CENTRAL TRAJECTORY

A Dissertation

APPROVED FOR THE

SCHOOL OF INDUSTRIAL ENGINEERING

BY

Dr. T.B. TRAFALIS

Dr. P.S. PULAT

Dr. S. RAMAN

Dr. F.H. GRANT

Dr. S. LAKSHMIVARAHAN

Acknowledgements

I wish to express sincere thanks to the many professors, friends, and colleagues who have helped make this study possible by their patient assistance and moral support at the various stages of research and writing. A special expression of gratitude goes to Dr. T. B. Trafalis, the major professor, who provided guidance and patient support above and beyond the call of duty. He has been my advisor since beginning of my Ph. D program. His expertise and guidance were of great importance for this study.

I also recognize and thank my committee members Dr. P. S. Pulat, Dr. S. Raman, Dr. S. Lakshmivarahan and Dr. F. H. Grant for their involvement. Their guidance and enthusiastic support throughout the research and completion of this study is sincerely appreciated. I also thank them for everything I have learned attending the various courses they have offered. I also wish to thank Dr. A. Ravindran and A. B. Badiru for their involvement as previous members of my committee.

Finally, heart-felt gratitude goes to my better half and lovely wife, Belir, for her love,constant support, encouragement, and patient understanding. I would like to thank my son, Deniz, and my daughter, Selin for their love and patience. I also thank to my parents, Murat and Selver, for their love and prayers throughout my studies.

Contents

List of Figures

Abstract

Interior Point Methods (IPMs) are iterative algorithms for mathematical optimization problems that can be interpreted as path-following procedures. Given a starting solution, the iterative scheme generates a sequence of points that converge to the optimal solution of the problem. The points specified by the iterative scheme lie on the central trajectory. The central trajectory is a smooth analytical curve in the interior of the feasible region of the problem. It starts from an interior point and ends at the optimal solution of the problem. Primal dual IPMs generate points that lie in the neighborhood of the central trajectory. The key ingredient of primal dual IPMs is the parameterization of the central trajectory. The duality gap depends linearly on the barrier parameter for the points in the central trajectory. In this research, a new approach to the parameterization of the central trajectory for primal dual IPMs is proposed. A continuous dynamical system that describes the rate of the change of the barrier parameter at the central trajectory is considered. Instead of parameterizing the central trajectory by the barrier parameter, it is parameterized by the time by describing a continuous dynamical system. Specifically, a new update rule based on the solution of an ordinary differential equation (ODE) for the barrier parameter

of the primal dual IPMs is presented. The resulting ordinary differential equation combined with the first order Karush-Kuhn-Tucker conditions, which are algebraic equations, are called differential algebraic equations (DAEs). By solving DAEs, we follow approximately the central trajectory of the primal dual IPMs. By doing so, we find an optimal solution to the given problem.

The proposed parameterization of the central trajectory for primal dual IPMs is investigated both for linear and convex quadratic optimization problems and primal dual IPMs are modified by using new parameterization. In addition, convergence, implementation, computational complexity and stability issues of the proposed parameterization of the central trajectory are also investigated. Some computational results for the proposed parameterization of the central trajectory for linear and convex quadratic optimization problems and applications to support vector machines (SVMs) for the classification problem are presented.

Chapter 1

Introduction

1.1 Overview

There has been an enormous research on using Interior Point Methods (IPMs) algorithms to solve optimization problems since Karmarkar's [37] study. IPMs are iterative algorithms for mathematical optimization problems that can be interpreted as trajectory-following procedures. In iterative algorithms, for a given starting solution, the iterative scheme generates a sequence of points that converge to the optimal solution of the problem. The points specified by the iterative scheme lie on a smooth analytical curve that is called *central trajectory*. A central trajectory starts from an interior point of the feasible region and ends at the optimal solution of the problem.

Primal dual IPMs, are considered the most successful techniques to solve linear optimization problems [40, 41]. Specifically, primal dual IPMs generate points that lie in the neighborhood of the central trajectory. The duality gap of the primal dual IPMs depends linearly on the barrier parameter for the points in the central trajectory and the convergence of the primal dual IPMs is achieved when the duality gap becomes close to zero. Therefore, the updating rule for the barrier parameter becomes so critical for the convergence of the primal dual IPMs. In brief, the key ingredient of primal dual IPMs is the parameterization of the central trajectory. Currently, the central trajectory is parameterized by the barrier parameter.

1.2 Research Objectives

In this research, a new approach to the parameterization of the central trajectory for primal dual IPMs is proposed. A continuous dynamical system that describes the rate of the change of the barrier parameter of the central trajectory is considered. Instead of parameterizing the central trajectory by the barrier parameter, we parameterize it by the time and therefore describe it by a continuous dynamical system. Specifically, a new update rule based on the solution of an ordinary differential equation (ODE) for the barrier parameter of the primal dual IPMs is presented.

First, we transform the given constrained optimization problem into an unconstrained optimization problem by adding barrier terms, corresponding to the constraints, with barrier parameters. Then, the unconstrained optimization problem is solved by using Newton's method. Rather than updating the barrier parameter by heuristic rules, we determine the developing trajectory of the barrier parameter by solving an ordinary differential equation. The resulting ordinary differential equation combined with the first order Karush-Kuhn-Tucker conditions, which are algebraic equations, are called differential algebraic equations (DAEs). By solving DAEs, we follow approximately the central trajectory of the primal dual IPMs. By doing so, we find an optimal solution to the given problem.

The proposed parameterization of the central trajectory for primal dual IPMs is investigated both for linear and convex quadratic optimization problems and primal dual IPMs are modified by using new parameterization. In addition, convergence, implementation, computational complexity and stability issues of the proposed parameterization of the central trajectory are also investigated.

1.3 Scope of the Dissertation

In Chapter 2 literature reviews of IPMs, DAEs and the central trajectory are given. Chapters 3, and 4, provide brief introductions to the primal dual IPMs for linear and convex quadratic optimization problems and DAEs respectively. Chapter 5 presents the proposed parameterization of the central trajectory of primal dual IPMs and its extensions to linear and convex quadratic optimization problems. Chapter 6 investigates convergence, implementation, computational complexity and stability issues of the proposed parameterization of the central trajectory. Chapter 7 reports computational results for some test problems. Chapter 8 concludes the dissertation and proposes some future research.

Through this dissertation, we use the name *optimization* instead of *programming*. In other words, the name *mathematical optimization* instead of *mathematical programming* or *linear optimization* (LO) instead of *linear programming* (LP) are used.

The main reason of using this terminology is to prevent the confusion that comes from the word *"programming"*. The word *"programming"* generally refers to the activity of writing computer programs. In some areas like combinatorial optimization and discrete optimization this terminology has already become generally accepted.

Chapter 2

Literature Review

2.1 Interior Point Methods

The methods for solving mathematical optimization problems can be categorized into two classes as Interior Point Methods (IPMs) and Exterior Point Methods respectively. Since Karmarkar's [37] revolutionary IPM for linear optimization (LO) problems, there has been a significant interest in studying optimization problems by using IPMs. Although IPMs have had great attention after Karmarkar's study in 1984, they had been introduced to the optimization field in the 1960s. In fact, some IPMs for nonlinear optimization problems, like the logarithmic barrier method in 1968 by Fiacco and McCormick [24], the affine scaling in 1967 by Dikin [19], and the analytical centers in 1967 by Huard [35] were studied. Although IPMs are not new, they are still under development. Many developments have been made on IPMs. These developments have resulted in a wide variety of IPMs. Various methods for

LO problems can be divided into the following three categories.

1. Projective methods: They are developed by Karmarkar [37], then studied in [2, 26, 61]. They use a potential function to measure the progress of each iteration. They have a polynomial-time convergence.

2. Affine scaling methods: They are originally developed by Dikin [19] and subsequently rediscovered in [4, 68]. The polynomial-time status of affine scaling methods is still not proven, but they work well in practice.

3. Path following methods and primal dual IPMs: They are based on the logarithmic barrier method. They are introduced by Renegar [56] and extended to the primal dual setting by Monteiro and Adler [48], and Kojima et al. [38]. They have a polynomial-time convergence.

Primal dual IPMs, also known as path following methods, are the most successful ones in practice. They are developed by using the following three general optimization methods:

1. Logarithmic barrier method: It is used to deal with inequality constraints.

2. Lagrange method: It is used to solve optimization problems with equality constraints.

3. Newton's method: It is used to solve nonlinear systems.

The logarithmic barrier method was originally introduced by Frisch [25] for convex optimization problems. Later, Fiacco and McCormick [24] developed the logarithmic barrier method for the following general inequality constrained optimization problems

$$\begin{aligned} \min \quad & f(x) \\ \text{subject to} \quad & g_i(x) \geq 0, \quad i = 1, \ldots, m. \end{aligned} \tag{2.1}$$

The method transforms this problem into an unconstrained optimization problem or into a sequence of unconstrained problems by adding a logarithmic term and a barrier parameter to the objective function as follows

$$\min \quad C(x, \mu) = f(x) - \mu \sum_{i=1}^{m} \log g_i(x). \tag{2.2}$$

We can see that, problem 2.2 is an unconstrained optimization problem for a given μ. $C(x, \mu)$ is called *logarithmic barrier function.* Because of the singularity of the logarithm at zero, the logarithmic barrier function will prevent the solution from going outside of the feasible region. Therefore, the logarithmic barrier method is considered as an IPM. The method starts from the interior of the feasible region, moves in a trajectory that generates a smooth curve in the interior of the feasible region and ends in the optimal solution. The barrier parameter μ is a positive number that monotonically decreases at each iteration. As it goes to zero, the objective function of the unconstrained optimization problem 2.2 becomes the objective function of the constrained problem 2.1 and the optimization problem 2.1 is solved.

The introduction of the new interior point algorithm by Karmarkar led researchers to reconsider the application of the logarithmic barrier method for linear and nonlinear optimization problems. Kojima et al. [38] and Monteiro and Adler [48] used the logarithmic barrier framework to present a new IPM that is named primal dual IPM.

They derive the first order optimality conditions associated with the minimization of a logarithmic barrier function, and define near solutions of this parameterized system of nonlinear equations. Some researchers like Nash et al. [50] and Breitfeld and Shanno [10] have successfully applied the logarithmic barrier method to large-scale optimization problems. Tutunji and Trafalis [63, 64, 66] have explored the logarithmic barrier methods with special emphasis in supervised neural network training.

Some researchers have proposed modifications on the logarithmic term. Polyak [54] proposed the following modified logarithmic barrier function by changing the logarithmic term. Specifically,

$$M(x,\mu,\lambda) = f(x) - \mu \sum_{i=1}^{m} \lambda_i \log\left(1 + \frac{g_i(x)}{\mu}\right), \tag{2.3}$$

where $\lambda = (\lambda_1, ..., \lambda_m)$ is a positive vector of Lagrange multipliers. Tuncel and Todd [65] have done some modifications in the logarithmic term in primal dual IPMs. They proposed to use the following entropy logarithmic term that has both primal and dual information:

$$\sum_{i=1}^{n} x_i s_i \ln(x_i s_i), \tag{2.4}$$

where x_j and s_j are the primal and dual variables respectively. Breitfeld and Shanno [9, 11] have also done some modifications on the logarithmic term.

Researchers [38, 48, 59, 43] have used some heuristics on the reduction of the barrier parameter μ. Mostly, the barrier parameter μ is reduced by a positive factor at each iteration. A vector of barrier parameters by associating a different barrier parameter μ_i to each constraint was presented in [18]. Predictor-corrector algorithm, a path following IPM, alternates between two type of steps: predictor steps, which improve the value of μ but which also tend to worsen the centrality measure, and corrector steps, which have no effect on the duality measure μ but improve centrality. The term predictor-corrector arose because of the analogy with predictor -corrector algorithms in ordinary differential equations (ODE). These algorithms follow the solution trajectory of an initial value ODE problem by alternating between predictor steps which move along a tangent to the trajectory and corrector steps which move back toward the trajectory from the predicted point [76]. In this study, we present an update rule based on an analytical update of the barrier parameter for the IPMs in a deterministic sense. A preliminary investigation and computational results of our study are presented in [62].

After great success of the IPMs on large-scale LO problems, many researchers have devoted their efforts to developing IPMs for nonlinear optimization problems. An extension of Karmarkar's algorithm and the trust region methods for quadratic optimization was studied in [78]. An IPM for solving smooth convex programs based on Newton's method was presented in [47]. Since the primal dual IPMs are considered the most effective techniques for solving large-scale LO problems, there have been some studies applying primal dual IPMs to more general classes of problems. Monteiro and Adler [49] have presented a primal dual IPM for convex quadratic optimization problems in [49]. After successful extensions of IPMs to convex and quadratic optimization problems, researchers in the IPM field are now trying to extend IPMs to more complex problems: as general nonlinear nonconvex optimization problems. The primal dual framework has been extended to general nonlinear optimization problems in [20].

In another study [69], an IPM for general nonlinear optimization has been described. Primal dual techniques for nonlinear optimization and their application to artificial neural network training have been studied in [17]. For an authoritative monograph on IPMs read the book by Nesterov and Nemirovskii [51]. For more recent developments, visit the online website http://www-unix.mcs.anl.gov/otc/InteriorPoint/ maintained by Nathan Brixius and Steve Wright at Argonne National Laboratory.

2.2 Central Trajectory

The central trajectory is fundamental to the study of IPMs for LO and has been subject of an enormous volume of research. Primal dual IPMs generate points that lie in the neighborhood of the central trajectory of the problem. The central trajectory of a problem is a smooth analytical curve which starts at the analytical center of the feasible region of the problem and moves to the optimal solution of the problem. The behaviour of the central trajectories of IPM fields was fully investigated in [5, 6, 46]. Central trajectories to the optimal set in LO problems are studied in [45]. By characterizing the affine scaling trajectories as solutions of certain parameterized logarithmic barrier families of problems, the convergence and limiting behavior of the affine scaling trajectories under no nondegeneracy assumptions have been analyzed and an analysis of the primal dual trajectories was also given in [1]. Using a different approach, convergence behavior of central trajectories arising from the logarithmic barrier function in LO without nondegeneracy assumption has been investigated in [74]. Fang and Puthenpura [21] showed that the moving directions of affine scaling methods and primal dual methods are merely the Newton directions along different central trajectories that lead to the solution of the Karush-Kuhn-Tucker (KKT) conditions of the given LO problem. Nunez and Freund [52] showed that certain properties of solutions along the central trajectory of the linear program are inherently related to the condition measures. They presented lower and upper bounds on sizes of optimal solutions along the central trajectory, and on rates of changes of solutions along the central trajectory as either the barrier parameter or the data of the linear program is changed.

The primal dual IPM maintains primal and dual feasibility and iterates to reduce the duality gap. The duality gap depends linearly on the barrier parameter for the points in the central trajectory. The rate of change of the barrier parameter can be considered as a continuous dynamical system. Since the pioneering study of Karmarkar, there has been a considerable interest in studying dynamical systems which solve LO problems. Faybusovich [22] introduced a class of dynamical systems which evolve in the interior of a given solution space, and solved linear, fractional-linear and convex optimization problems. He described a parameterization of a given solution space by a smooth manifold to get rid of inequality constraints. Then, he showed that the solution of the corresponding gradient system converges to a local extremum of the initial optimization problem. In another study, [23] he introduced a new class of completely integrable Hamiltonian systems that solve LO problems.

More recently primal dual IPMs have been applied successfully in semidefinite optimization. the book by Wolkowicz, Saigal, and Vanderberghe [75] provides a collection of review papers and a large number of references. For more recent developments in semidefinite optimization visit the homepage http://www-user.tu-chemnitz.de/ helmberg/semidef.html by Helmberg.

2.3 Differential Algebraic Equations

The description of some physical problems by differential equations involves a nonlinear system or algebraic equations, too. A differential equation and a nonlinear system form a Differential-Algebraic Equations (DAE) system. The DAE system occurs frequently as an initial value problem in modeling electrical networks, the flow of incompressible fluids, mechanical systems subject to constraints, robotics, distillation process, power systems, and in many other applications [12, 33, 39].

In general, DAE system can be expressed in the following semi-explicit form

$$\begin{aligned} \frac{dx}{dt} = x' &= f(x,y,t) \\ 0 &= g(x,y,t), \end{aligned} \tag{2.5}$$

where $x(t)$ and $y(t)$ are algebraic variables. DAE systems are different from ordinary differential equation (ODE) systems. Here the ODE for $x(t)$ depends on additional algebraic variables $y(t)$ and the solution is forced to satisfy the nonlinear system. In the case of a DAE, the algebraic equations help to determine the solution.

In recent years DAEs have attracted much interest - partially because of their importance as models for a large class of dynamical processes, e.g. in mechanics, robotics, chemical engineering but also because of their intrinsic numerical difficulties.

In many cases DAE system can be solved efficiently by means of standard numerical methods for ODE systems. This approach appears to have been introduced by Gear [27], and since then it has been used by many researchers [12, 33, 34, 39, 42, 53]. Gear and Petzold [28] studied ODE methods for the solution of DAE systems. The numerical solutions of DAEs by Runge-Kutta methods were studied by Hairer et al. [31]. The numerical solutions of DAEs of index one and higher indices by multi-step methods, Runge-Kutta methods, Rosenbrock method and extrapolation methods were presented by Hairer and Wanner [32]. Some new algorithms and software for sensitivity analysis of DAEs were presented by Maly and Petzold [42]. They showed that the proposed algorithms are well suited for large-scale problems. Algorithms and relevant software packages that are available for ODEs and DAEs were described by Hindmarsh and Petzold [33, 34]. A number of difficulties which can arise when numerical methods are used to solve DAE systems have been outlined by Petzold [53]. Rheinboldt [58] proved an existence and uniqueness theory for DAEs based on the theory of ODEs on a manifold.

Over the last two decades, methods based on ODEs for solving optimization problems had some attention in parallel to the developments of IPMs. Brown and Bartholomew-Biggs [13] have shown, by means of numerical experiments, that unconstrained optimization techniques based on the solution of the system of ODE can compare very favorably with some classical optimization techniques as regard reliability, accuracy, and efficiency.

During the last decade, the optimization of systems involving differential-algebraic equations has become an important research area in applications involving chemical processing, robotics, structural analysis, and aerospace engineering. DAEs are connected to optimization in at least two ways. On the one hand variational principles are used to formulate DAE, e.g. in multibody dynamics or in boundary value problems associated with optimal control problems. On the other hand many relevant practical problems, which are modeled by DAE, call for optimization rather than forward simulation alone. Xiong et al. [77] presented a differential-algebraic approach to linear programming. Gopal and Biegler [30] presented a successive linear programming (SLP) approach for the solution of the DAE derivative array equations for the initialization problem. Renfro [57] did computational studies in the optimization of systems described by DAEs. In another study [15], a novel nonlinear programming strategy by using IPMs is developed and applied to the optimization of DAE systems.

Chapter 3

Primal Dual Interior Point Methods

3.1 Linear Optimization Problems

Next we describe briefly the primal dual IPM for the linear optimization (LO) problems. Our description is based in references [43, 44, 48, 76].

Let us consider the primal LO problem P in standard form

$$
\begin{aligned}
P: \qquad & \min \quad f(x) = c^T x \\
& \text{subject to} \\
& \qquad\qquad Ax = b \\
& \qquad\qquad x \geq 0,
\end{aligned}
\tag{3.1}
$$

where $x \in \Re^n$ is a vector of decision variables, $A \in \Re^{m \times n}$ is a coefficient matrix of constraints, $c \in \Re^n$ is a coefficient vector of the objective function $f(x)$, and $b \in \Re^m$ is a right hand side (RHS) vector of constraints.

The dual problem D of the problem P becomes

$$
\begin{aligned}
D: \qquad & \max \quad g(y) = b^T y \\
& \text{subject to} && (3.2) \\
& \qquad A^T y + z = c \\
& \qquad\qquad z \geq 0,
\end{aligned}
$$

where $y \in \Re^m$ is a vector of dual variables and $z \in \Re^n$ is a vector of dual slack variables.

It is assumed that A has a full row-rank. In other words $m \leq n$ and the feasible regions of the problems P and D are not empty.

To solve either problem P or D, we transform it into an unconstrained optimization problem. First of all, inequality constraints are to be dealt with. Inequality constraints of P and D can be handled by the logarithmic barrier method by adding a logarithmic barrier term to the objective function of P and D. Adding a barrier term to the objective functions of P and D results to the following problems P_μ and D_μ respectively.

$$
\begin{aligned}
& \min \quad f(x,\mu) = c^T x - \mu \sum_{i=1}^{n} \ln x_i \\
P_\mu: \qquad & \text{subject to} && (3.3) \\
& \qquad Ax = b \\
& \qquad \mu > 0
\end{aligned}
$$

$$
D_\mu: \quad \begin{aligned} \max \quad g(y,z,\mu) &= b^T y - \mu \sum_{i=1}^{n} \ln z_i \\ \text{subject to} & \\ A^T y + z &= c \\ \mu &> 0. \end{aligned} \tag{3.4}
$$

By adding a barrier term, in addition to minimizing an objective function, moving away from the boundary of the feasible region is also achieved. The barrier parameter $\mu > 0$ is a balance parameter of two conflicting actions of minimizing an objective function and moving away from boundary of the feasible region. Some $\mu > 0$ can be chosen to start and then let it approach to 0 during the iteration process.

After dealing with inequality constraints, equality constraints are to be dealt with. Equality constraints of P_μ and D_μ can be handled by Lagrange's Method [8]. To solve P_μ and D_μ for a given barrier parameter $\mu > 0$, corresponding Lagrangian functions can be constructed as follows.

$$
L_P(x,y,\mu) = c^T x - \mu \sum_{i=1}^{n} \ln x_i - y^T(Ax - b) \tag{3.5}
$$

$$
L_D(x,y,z,\mu) = b^T y - \mu \sum_{i=1}^{n} \ln z_i - x^T(A^T y + z - c). \tag{3.6}
$$

Note that y in equation 3.5 and x in equation 3.6 correspond to Lagrangian multipliers but they are also dual variables of each problem because of Lagrangian multipliers' interpretation as dual variables.

Now, an unconstrained optimization problem 3.5 or 3.6 has been obtained. To find an optimal solution to the unconstrained optimization problem, Karush-Kuhn-Tucker (KKT) conditions need to be satisfied. Let e be the column vector of all

$1's$ and X and Z be $n \times n$ diagonal matrices defined by $X = diag(x_1, \ldots, x_n)$ and $Z = diag(z_1, \ldots, z_n)$ respectively. Then, the first order KKT necessary conditions for minimizing $L_P(x, y, \mu)$ are given by:

$$\begin{aligned} c - \mu X^{-1}e - A^T y &= 0 \\ Ax - b &= 0. \end{aligned} \tag{3.7}$$

The first order KKT necessary conditions for maximizing $L_D(x, y, z, \mu)$ are given by:

$$\begin{aligned} A^T y + z &= c \\ Ax - b &= 0 \\ \mu Z^{-1}e - x &= 0. \end{aligned} \tag{3.8}$$

In fact, the KKT necessary conditions 3.7 and 3.8 characterizing the optimum of the problem P and D can be combined and re-formulated as follows.

$$\begin{aligned} Ax - b &= 0 \\ A^T y + z - c &= 0 \\ XZe - \mu e &= 0. \end{aligned} \tag{3.9}$$

In this formulation, the first equation maintains primal feasibility, the second equation maintains dual feasibility and the last equation maintains an approximation to complementary slackness condition. Note that $\mu = 0$ corresponds to the ordinary complementary slackness condition.

The KKT necessary conditions presented in 3.9 construct a nonlinear system in the form of $F(x, y, z) = 0$. Newton's method is one of the most commonly used techniques to find a root of a nonlinear system by iteratively approximating the system by linear equations. In other words, we need to find x^k, y^k,and z^k such that $F(x^k, y^k, z^k) = 0$. By using Taylor's approximation at $x = x^k$, $y = y^k$, $z = z^k$ we obtain a linear approximation

$$F(x^k + dx, y^k + dy, z^k + dz) \approx F(x^k, y^k, z^k) + J(x^k, y^k, z^k)(dx, dy, dz)^T, \tag{3.10}$$

where $J(x^k, y^k, z^k)$ is the Jacobian of $F(x^k, y^k, z^k)$ and dx, dy, dz are the moving direction vectors. Since the left hand side of 3.10 evaluates at a root of $F(x, y, z) = 0$, we have the following linear system

$$F(x^k, y^k, z^k) = -J(x^k, y^k, z^k)(dx, dy, dz)^T. \tag{3.11}$$

A solution of equation 3.11 provides one Newton iteration from x^k, y^k, z^k to $x^{k+1} = x^k + dx$, $y^{k+1} = y^k + dy$, $z^{k+1} = z^k + dz$ with dx, dy, dz moving direction vectors with a unit step length.

Now, let us focus on the nonlinear system 3.9. The Jacobian of the nonlinear system 3.9 is equal to the following matrix

$$J(x^k, y^k, z^k) = \begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Z & 0 & X \end{bmatrix}. \tag{3.12}$$

By setting some $\mu > 0$, the vectors x^k, y^k, z^k that solve the nonlinear system 3.9 are obtained by using Newton's method. Since these vectors are dependent on the choice of the barrier parameter μ, we get a family of solutions depending on the value of μ. The *central trajectory* is defined as the set of all vectors $x^k(\mu), y^k(\mu)$ and $z^k(\mu)$, satisfying the nonlinear system 3.9. The limits of $x(\mu), y(\mu)$ and $z(\mu)$ as μ goes to zero approach to the solution.

Given $x^0 > 0, y^0$ and $z^0 > 0$ and $\mu > 0$, the moving direction vectors dx, dy, and dz that move from the current iterate to a new iterate while satisfying the KKT necessary conditions are determined by using the following linearized system as defined in 3.11

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = - \begin{bmatrix} Ax^0 - b \\ A^T y^0 + z^0 - c \\ \mu e - X^0 Z^0 e \end{bmatrix} = \begin{bmatrix} d_P \\ d_D \\ d_w \end{bmatrix}. \tag{3.13}$$

The solution of the linear system 3.13 with respect to the moving direction vectors dx, dy, and dz can be expressed as follows.

$$\begin{aligned} dy &= (AZ^{-1}XA^T)^{-1}(AZ^{-1}Xd_D + d_P - AZ^{-1}d_w) \\ dz &= d_D - A^T dy \\ dx &= Z^{-1}(d_w - Xdz). \end{aligned} \tag{3.14}$$

The new points for the next iteration for the P and D problems are given as follows:

$$\begin{aligned} x^1 &= x^0 + \rho\alpha_P dx \\ y^1 &= y^0 + \rho\alpha_D dy \\ z^1 &= z^0 + \rho\alpha_D dz, \end{aligned} \tag{3.15}$$

where $0 < \rho < 1$, and α_P and α_D are step sizes for the primal and dual problem respectively. Step sizes are chosen to assure that x and z are positive as follows

$$\begin{aligned} \alpha_P &= \min\left\{-\frac{x_i}{dx_i} : \forall\ dx_i < 0, \quad 1 \le i \le n\right\} \\ \alpha_D &= \min\left\{-\frac{z_i}{dz_i} : \forall\ dz_i < 0, \quad 1 \le i \le n\right\}. \end{aligned} \tag{3.16}$$

Convergence of the primal dual IPM has been achieved when the duality gap becomes close to zero with the given accuracy ε. Specifically,

$$c^T x - b^T y = \varepsilon. \tag{3.17}$$

Moreover, there exists a simple relation between the duality gap and the barrier parameter, μ. Specifically,

$$c^T x - b^T y = c^T x - (Ax)^T y = x^T(c - A^T y) = x^T z = n\mu. \tag{3.18}$$

We can see that the duality gap depends linearly on the barrier parameter for the points in the central trajectory. As we mentioned before, the barrier parameter μ starts with some $\mu > 0$ and then we let it approach zero by updating it at each iteration. Note that equation 3.18 suggests us the following update procedure for the barrier parameter μ

$$\mu^{k+1} = \sigma \frac{(x^k)^T z^k}{n} = \sigma \mu^k, \quad \text{where} 0 < \sigma < 1. \tag{3.19}$$

This update rule of the barrier parameter μ is mostly used in the primal dual IPM studies.

The primal dual IPM continues to iterate by reducing μ as in equation 3.19 and computing the new moving direction vectors by using equation 3.14 and new points by using equation 3.15 until convergence is achieved. An algorithm for the primal dual IPM for LO problems can be stated as follows.

Primal Dual IPM Algorithm for LO Problems

Step 0 Initialization:

- SET iteration number $k = 0, \varepsilon, \rho$ and σ.
- START with an initial solution $x^k > 0, y^k, z^k > 0$ and $\mu^k > 0$.

Step 1 Stopping Criteria:

- IF equation 3.17 is satisfied GO TO Step 3.
- ELSE GO TO Step 2.

Step 2 Iteration:

- COMPUTE moving directions dx^k, dy^k, dz^k, by equation 3.14.
- COMPUTE step sizes α_P and α_D, by equation 3.16.
- MOVE to the next solution $x^{k+1}, y^{k+1}, z^{k+1}$, using equation 3.15.
- UPDATE the barrier parameter μ^k, by equation 3.19.
- SET $k = k + 1$.
- GO TO Step 1.

Step 3 Stop with the optimal point $x^*, y^*, z^* = x^k, y^k, z^k$.

3.2 Convex Quadratic Optimization Problems

Next we describe briefly the primal dual IPM for the convex quadratic optimization (QO) problems as presented in references [18, 21, 48].

Let us consider the primal QO problem QP in its standard form

$$
QP: \quad \begin{aligned}
\min \quad f(x) &= c^T x + \frac{1}{2} x^T Q x \\
\text{subject to} & \\
Ax &= b \\
x &\geq 0,
\end{aligned} \tag{3.20}
$$

where $x \in \Re^n$ is a vector of decision variables, $A \in \Re^{m \times n}$ is a matrix of constraints, $Q \in \Re^{n \times n}$ is a symmetric positive semi definite coefficient matrix of the quadratic terms, $c \in \Re^n$ is a coefficient vector of the linear terms of the objective function $f(x)$, and $b \in \Re^m$ is a right hand side (RHS) vector of constraints.

The dual problem QD of the problem QP becomes

$$
QD: \quad \begin{aligned}
\max \quad g(y, x) &= b^T y - \frac{1}{2} x^T Q x \\
\text{subject to} & \\
A^T y - Qx + z &= c \\
z &\geq 0,
\end{aligned} \tag{3.21}
$$

where $y \in \Re^m$ is a vector of dual variables and $z \in \Re^n$ is a vector of dual slack variables.

It is considered that A has a full row-rank, $m \leq n$ and the feasible regions of the problems QP and QD are not empty.

To solve either problem QP or QD, we transform it into an unconstrained optimization problem. First of all, inequality constraints will be dealt with as explained in Section 3.1 by adding a logarithmic barrier term to the objective function of QP and QD. Adding a barrier term to the objective functions of QP and QD results to the following problem QP_μ and QD_μ respectively.

$$QP_\mu: \quad \begin{array}{rcl} \min \quad f(x,\mu) & = & c^T x + \frac{1}{2} x^T Q x - \mu \sum_{i=1}^{n} \ln x_i \\ \text{subject to} & & \\ Ax & = & b \\ \mu & > & 0 \end{array} \tag{3.22}$$

$$QD_\mu: \quad \begin{array}{rcl} \max \quad g(y,x,z,\mu) & = & b^T y - \frac{1}{2} x^T Q x - \mu \sum_{i=1}^{n} \ln z_i \\ \text{subject to} & & \\ A^T y - Qx + z & = & c \\ \mu & > & 0. \end{array} \tag{3.23}$$

Equality constraints of QP_μ and QD_μ can be handled by Lagrange's Method [8]. To solve QP_μ and QD_μ for a given barrier parameter $\mu > 0$, corresponding Lagrangian functions can be constructed as follows

$$L_{QP}(x,y,\mu) = c^T x + \frac{1}{2} x^T Q x - \mu \sum_{i=1}^{n} \ln x_i - y^T (Ax - b), \tag{3.24}$$

$$L_{QD}(x,y,z,\mu) = b^T y - \frac{1}{2} x^T Q x - \mu \sum_{i=1}^{n} \ln z_i - x^T (A^T y - Qx + z - c). \tag{3.25}$$

The associated KKT necessary conditions referring to the optimization of $L_{QP}(x, y, \mu)$ and $L_{QD}(x, y, z, \mu)$ become the following system

$$
\begin{aligned}
Ax - b &= 0 \\
-Qx + A^T y + z - c &= 0 \\
XZe - \mu e &= 0.
\end{aligned} \tag{3.26}
$$

The KKT necessary conditions presented in 3.26 is a nonlinear system in the form of $F(x, y, z) = 0$. The Jacobian of the nonlinear system 3.26 is assumed to be nonsingular and it is equal to the following matrix

$$
J(x^k, y^k, z^k) = \begin{bmatrix} A & 0 & 0 \\ -Q & A^T & I \\ Z & 0 & X \end{bmatrix}. \tag{3.27}
$$

By setting some $\mu > 0, x^k, y^k, z^k$ that solves the nonlinear system 3.26 will be obtained by using Newton's method. Given $x^0 > 0, y^0$ and $z^0 > 0$ and $\mu > 0$, moving direction vectors dx, dy, and dz that move from the current iterate to a new iterate while satisfying the KKT necessary conditions can be found. Moving direction vectors dx, dy, and dz are determined by the following system of linear equations

$$
\begin{bmatrix} A & 0 & 0 \\ -Q & A^T & I \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = - \begin{bmatrix} Ax^0 - b \\ -Qx^0 + A^T y^0 + z^0 - c \\ \mu e - X^0 Z^0 e \end{bmatrix} = \begin{bmatrix} d_P \\ d_D \\ d_w \end{bmatrix}. \tag{3.28}
$$

The solution of the linear system 3.28 with respect to the moving direction vectors dx, dy, and dz can be expressed as follows

$$
dy = [A(Z + XQ)^{-1} X A^T]^{-1} [A(Z + XQ)^{-1} X d_D + d_P - A(Z + XQ)^{-1} d_w],
$$

$$
\begin{aligned}
dx &= (Z + XQ)^{-1}(d_w - Xd_D + XA^T dy), \\
dz &= X^{-1}(d_w - Zdx).
\end{aligned} \tag{3.29}
$$

The new points for the next iteration for the QP and QD problems are given as

$$
\begin{aligned}
x^1 &= x^0 + \rho\alpha_P dx \\
y^1 &= y^0 + \rho\alpha_D dy \\
z^1 &= z^0 + \rho\alpha_D dz,
\end{aligned} \tag{3.30}
$$

where $0 < \rho < 1$, and α_P and α_D are step sizes for the primal and dual problem respectively. Step sizes are chosen to assure that x and z are positive as follows

$$
\begin{aligned}
\alpha_P &= \min\left\{-\frac{x_i}{dx_i} : \forall\ dx_i < 0, \quad 1 \le i \le n\right\} \\
\alpha_D &= \min\left\{-\frac{z_i}{dz_i} : \forall\ dz_i < 0, \quad 1 \le i \le n\right\}.
\end{aligned} \tag{3.31}
$$

Convergence of the primal dual IPMs for QO problems has also been achieved when the duality gap becomes close to zero or the given accuracy ε. Specifically,

$$
c^T x - b^T y + x^T Qx = \varepsilon. \tag{3.32}
$$

Again there exists a simple relation between the duality gap and the barrier parameter, μ for the QO problems

$$
c^T x - b^T y + x^T Qx = c^T x - (Ax)^T y + x^T Qx = x^T(c - A^T y + Qx) = x^T z = n\mu. \tag{3.33}
$$

We see that the duality gap also depends linearly on the barrier parameter for the points in the central trajectory for the QO problems. As we mentioned before, the barrier parameter μ starts with some $\mu > 0$ and we let it approach zero by updating it at each iteration. Note that, equation 3.33 suggests us the following update procedure for the barrier parameter μ.

$$\mu^{k+1} = \sigma \frac{(x^k)^T z^k}{n} = \sigma \mu^k, \quad \text{where} 0 < \sigma < 1. \tag{3.34}$$

The primal dual IPM for QO problems continues to iterate by reducing μ as in equation 3.34 and computing the new moving direction vectors by using equation 3.29 and new points by using equation 3.30 until convergence is achieved. An algorithm for the primal dual IPM for QO problems can be stated as follows.

Primal Dual IPM Algorithm for QO Problems

Step 0 Initialization:

- SET iteration number $k = 0, \varepsilon, \rho$ and σ.
- START with an initial solution $x^k > 0, y^k, z^k > 0$ and $\mu^k > 0$.

Step 1 Stopping Criteria:

- IF equation 3.32 is satisfied GO TO Step 3.
- ELSE GO TO Step 2.

Step 2 Iteration:

- COMPUTE moving directions dx^k, dy^k, dz^k, by equation 3.29.
- COMPUTE step sizes α_P and α_D, using equation 3.31.
- MOVE to the next solution $x^{k+1}, y^{k+1}, z^{k+1}$, by equation 3.30.
- UPDATE the barrier parameter μ^k, as in equation 3.34.
- SET $k = k + 1$.
- GO TO Step 1.

Step 3 Stop with the optimal point $x^*, y^*, z^* = x^k, y^k, z^k$.

Chapter 4

Differential Algebraic Equations (DAEs)

4.1 Introduction to DAEs

Mathematical models of some engineering, physical, and scientific problems frequently take the following explicit form of a system of ordinary differential equations (ODEs)

$$\frac{dx}{dt} = x' = f(x, t), \tag{4.1}$$

where, t is time and x is a vector of dependent variables or state variables. The *initial value problem* for the equation (4.1) is to find the solution $x(t)$ that satisfies a given initial condition $x(t_0) = x_0$.

In some cases, the model also involves dependent variables whose time derivatives do not appear in the equations. In other words, equation (4.1) also involves algebraic equations too. The set of equations which is the combination of both differential and algebraic equations that defines this model is known as a *differential-algebraic equation* (DAE) system. The most general DAE system is expressed in the fully implicit form as

$$F(x, x', t) = 0, \tag{4.2}$$

where F is some function. Another way to present a DAE system is to use the following semi-explicit form

$$\begin{aligned} \frac{dx}{dt} = x' &= f(x, y, t) \\ 0 &= g(x, y, t), \end{aligned} \tag{4.3}$$

where y is another vector of dependent variables. Note that y is a variable of the defining system of ODE (4.1) but dy/dt does not appear in the system. The DAE systems occurs frequently as an initial value problem in modeling electrical networks, the flow of incompressible fluids, mechanical systems subject to constraints, robotics, distillation process, power systems, and in many other applications [12, 33, 39].

DAE systems are more general than ODE systems.They include ODE systems as a special case, but they also include problems that are quite different from ODEs in nature [53]. In the case of a DAE, the algebraic equation helps to determine the solution. The differences and similarities between DAEs and ODEs can be summarized as follows: an ODE involves integration. On the other hand, a DAE involves both integrations and differentiations. Since DAEs involve both integrations and differentiations, one may hope that performing analytical differentiations to a given system and eliminating as needed will result in an explicit ODE for all unknowns. This turns out to be true unless the problem is singular.

4.2 Index of DAEs

Since a DAE involves both integrations and differentiations, by applying analytical differentiations to a given system and eliminating as needed will yield an explicit ODE system. A property known as the *index* plays a key role in the classification and behavior of DAEs. *Index* is defined as the minimum number of times that all or part of the DAE system must be differentiated to get a system of ODEs. Let us consider the DAE system in equation 4.3. If we differentiate the algebraic equation of 4.3 with respect to t, we will get the following:

$$\begin{aligned} \frac{dx}{dt} &= f(x,y,t) \\ g_x(x,y,t)\frac{dx}{dt} + g_y(x,y,t)\frac{dy}{dt} &= -g_t(x,y,t), \end{aligned} \tag{4.4}$$

or equivalently

$$\begin{aligned} \frac{dx}{dt} &= f(x,y,t) \\ g_y(x,y,t)\frac{dy}{dt} &= -g_t(x,y,t) - g_x(x,y,t)f(x,y,t). \end{aligned} \tag{4.5}$$

If g_y is nonsingular, then equation 4.5 is an implicit ODE and we say that DAE in equation 4.3 has index one. If this is not the case, suppose that with algebraic manipulation and coordinate changes we can rewrite equation 4.4 in the form of the algebraic system of equation 4.3 but with different x and y. Again we differentiate the algebraic equation. If an implicit ODE results, we say that the original problem has index two. If the new system is not an implicit ODE, we repeat the process. The number of differentiation steps required in this procedure is known as the *index* of DAE system in 4.3.

ODE systems in equation 4.1 have index zero. The DAEs in equation 4.3 have index one if g_y is nonsingular. In a sense, the index of a DAE system is a measure of its degree of singularity. The more singular a DAE system is, the more difficult to solve it numerically. Therefore, DAEs of index one are easy to solve.

4.3 Numerical Methods for DAEs

Semi-explicit DAEs of index one can be treated as ODEs. Let us consider the semi-explicit form of a DAE system as in 4.3. In this problem $y(t)$ can be determined for a given $x(t)$ by solving $g(x, y, t) = 0$. With suitable smoothness assumptions and an initial y_0 such that $g(x_0, y_0) = 0$, the existence of the inverse of the Jacobian g_y guarantees that y can be written as a function of x, namely $y(t) = G(x(t))$. Then $x' = f(x, G(x), t)$, where x_0 is given, is an initial value problem for $x(t)$. Thus, one way to solve index one problems is to apply an integrator to differential equation $x' = f(x, y, t)$ and every time that it needs to evaluate $f(x, y, t)$ for a specific x, solve algebraic equations $g(x, y, t) = 0$ for the corresponding y and then substitute x and y in $f(x, y, t)$. Then the resulting DAE system can be solved by using ODE methods.

The idea of using ODE methods for solving DAE systems directly was introduced by Gear [27]. The simplest way to solve DAEs of index 1 is to apply first order back differentiation formulas or backward Euler's method. In this method, the derivative $x'(t_n)$ is approximated by a backward difference of $x(t_{n-1})$, and the resulting system of nonlinear equations is solved by $x(t_n)$.

Let us consider the general form of a DAE system as in 4.2. Applying the backward

Euler's method to this system, we obtain

$$0 = F(x_n, \frac{x_n - x_{n-1}}{h}, t_n), \tag{4.6}$$

where h is the step size. This results in general to a system of nonlinear equations of x_n at each time step n. In this method the solution is advanced from time t_{n-1} to t_n. Higher order techniques such as backward differentiation formulas, Runge-Kutta methods, and extrapolation methods are generalizations of this idea [3].

Now, let us consider the backward Euler method applied to the semi-explicit form of a DAE system as in 4.3. Specifically,

$$\begin{aligned} \frac{x_n - x_{n-1}}{h_n} &= f(x_n, y_n, t_n) \\ 0 &= g(x_n, y_n, t_n), \end{aligned} \tag{4.7}$$

where h is the step size. Solving algebraic equations for y_n will result into $y_n = G(x_n, t_n)$ and substituting into difference equation (4.7) yields

$$\frac{x_n - x_{n-1}}{h_n} = f(x_n, G(x_n), t_n), \tag{4.8}$$

which is just the backward Euler approximation of the corresponding ODE.

One of the main advantages in using ODE methods directly for solving DAE systems is that ODEs preserve sparsity of the system. As mentioned before DAEs can be reduced to an ODE form by differentiating under a non-singularity assumption. In this case, we need to invert a matrix that may not result to a sparse matrix while the original one is sparse. This approach also tends to remove the natural sparsity of the system. The most challenging difficulties for solving DAE systems arise when the non-singularity assumption fails. DAE systems that are not of index one can not be solved so simply. As we see in the next chapter, while parameterizing the central trajectory, we end up with a differential equation and a nonlinear system that form a DAE system. Specifically, this DAE system has index one. Therefore, numerical methods for DAEs for higher indices were not considered in this dissertation.

Chapter 5

A New Approach to the Parameterization of the Central Trajectory

5.1 Introduction

The primal dual IPM maintains primal feasibility and dual feasibility and iterates to reduce the duality gap. The duality gap depends linearly on the barrier parameter for the points in the central trajectory. Our objective is to consider a continuous dynamical system that describes the rate of change of the barrier parameter. In this chapter, we develop a new approach to the parameterization of the central trajectory of the primal dual IPM for LO and convex QO problems respectively. First of all, given a constrained optimization problem will be transformed into an unconstrained optimization problem by adding barrier terms for the constraints with barrier parameters. Then, the unconstrained optimization problem will be solved iteratively by using Newton's method. Rather than updating the barrier parameter by heuristic rules, we have determined the developing trajectory of the barrier parameter by using an ordinary differential equation. The resulting differential equation combined with algebraic equations, that are first order Karush-Kuhn-Tucker (KKT) conditions, form a differential algebraic equation (DAE) system. The DAE is used to determine the central trajectory of the optimization problem. By doing so, we find an optimal solution to the given problem by following the central trajectory.

5.2 Current Parameterization of Central Trajectory

The key ingredient of primal dual IPMs is the parameterization of the central trajectory, since primal dual IPMs generate points that lie in the neighborhood of the central trajectory. As explained in chapter 3, the duality gap depends linearly on the barrier parameter for the points in the central trajectory. Therefore the current parameterization of the central trajectory is based on the barrier parameter μ. In chapter 3, the central trajectory is defined as the set of all vectors $x^k(\mu), y^k(\mu)$ and $z^k(\mu)$, satisfying the KKT conditions. Indeed, the KKT conditions are the necessary and sufficient conditions for $x(\mu), y(\mu)$ and $z(\mu)$, being on the central trajectory. The limits of $x(\mu), y(\mu)$ and $z(\mu)$, as μ goes to zero approach to the optimal solution. Since these vectors are dependent on the choice of the barrier parameter μ, we get a family of solutions depending on the value of μ.

In the rest of the chapter, we present a new approach to the parameterization of the central trajectory by considering a continuous dynamical system that describes the rate of change of the barrier parameter μ. Hence, we parameterize the central trajectory by time, t. As a result, the central trajectory will be defined as the set of all vectors $x^k(t), y^k(t), z^k(t)$ and $\mu(t)$ satisfying the KKT necessary conditions. As time increases, $\mu(t)$ will decrease to zero and $x(t), y(t)$, and $z(t)$ approach to the optimal solution.

5.3 New Parameterization of the Central Trajectory for LO Problems

Let us consider the primal LO problem P in the standard form

$$
\begin{array}{llrcl}
P: & \min & f(x) & = & c^T x \\
 & \text{subject to} & & & \\
 & & Ax & = & b \\
 & & x & \geq & 0,
\end{array}
\tag{5.1}
$$

where $x \in \Re^n$ is a vector of decision variables, $A \in \Re^{m \times n}$ is a coefficient matrix of constraints of full rank, $c \in \Re^n$ is a coefficient vector of the objective function $f(x)$, and $b \in \Re^m$ is a right hand side (RHS) vector of constraints.

The corresponding P_μ is as follows

$$
\min \quad f(x, \mu) \quad = \quad c^T x - \mu \sum_{i=1}^{n} \ln x_i
$$

$$P_\mu: \qquad \text{subject to} \tag{5.2}$$

$$\begin{aligned} Ax &= b \\ x &> 0 \end{aligned}$$

where $\mu > 0$ is the barrier parameter. The corresponding first order KKT necessary conditions are given as follows

$$\begin{aligned} b - Ax &= 0 \\ A^T y + z &= c \\ XZe - \mu e &= 0. \end{aligned} \tag{5.3}$$

The problem P_μ can also be defined as follows:

$$\theta(\mu) = \inf \left\{ c^T x - \mu \sum_{i=1}^{n} \ln x_i \quad \text{such that} \quad Ax = b, x > 0 \right\}. \tag{5.4}$$

Now, we state the following theorem by Bazaraa et al. [7] (Theorem 9.4.3) that will be used in our analysis.

Theorem 5.1 *Let $f : R^n \to R$, and $g : R^n \to R^m$ be continuous functions, and let X be a nonempty closed set in R^n. Suppose that the set $\{x \in X : g(x) < 0\}$ is not empty. Furthermore, suppose that the primal problem to minimize $f(x)$ subject to $g(x) \leq 0, x \in X$ has an optimal solution x^* with the following property. Given any neighborhood N around x^*, there exists an $x \in X \cap N$ such that $g(x) < 0$. Then,*

$$\textit{minimum}\,\{f(x) : g(x) \leq 0, x \in X\} = \lim_{\mu \to 0^+} \theta(\mu) = \inf_{\mu > 0} \theta(\mu).$$

Letting $\theta(\mu) = f(x_\mu) + \mu B(x_\mu)$, where $x_\mu \in X$ and $g(x_\mu) < 0$ and B is a barrier function that is nonnegative and continuous over the region $\{x : g(x) < 0\}$, then the *limit of any convergent subsequence of $\{x_\mu\}$ is an optimal solution to the primal problem and furthermore, $\mu B(x_\mu) \to 0$ as $\mu \to 0^+$.*

By using the above theorem, it is stated that the optimal solution to the problem P could be obtained by minimizing $\theta(\mu)$. In other words,

$$\text{minimum}\left\{c^T x \quad \text{such that} \quad Ax = b, x \geq 0\right\} = \inf_{\mu>0} \theta(\mu). \tag{5.5}$$

Note that, 5.1 and 5.4 are strictly convex optimization problems because both the objective function is strictly convex and the constraints are convex. Consequently, for any fixed $\mu, \theta(\mu)$ has a unique minimum. The minimum of $\theta(\mu)$ will be found by using the steepest descent method. Next, we need to consider μ as a function of parameter t. To find the rate of change of the barrier parameter μ, we have to move in the direction of the negative gradient of $\theta(\mu)$. Thus,

$$\frac{d\mu}{dt} = -\frac{d\theta(\mu)}{d\mu}. \tag{5.6}$$

From equation 5.4, $\theta(\mu)$ is differentiable. Now, let us find its derivative with respect to μ (note that x is a function of μ)

$$\frac{d\theta(\mu)}{d\mu} = c^T \frac{dx}{d\mu} - \sum_{i=1}^{n} \ln x_i - \mu \sum_{i=1}^{n} \frac{1}{x_i} \frac{dx_i}{d\mu} \quad \text{such that} \quad A\frac{dx}{d\mu} = 0. \tag{5.7}$$

If we consider the second equation from KKT necessary conditions 5.3, $c^T = y^T A + z^T$, and substitute in 5.7 then we have

$$\frac{d\theta(\mu)}{d\mu} = z^T \frac{dx}{d\mu} - \sum_{i=1}^{n} \ln x_i - \mu \sum_{i=1}^{n} \frac{1}{x_i} \frac{dx_i}{d\mu} \quad \text{such that} \quad A\frac{dx}{d\mu} = 0. \tag{5.8}$$

We can see that the first and the last terms cancel each other each other by the help of the third equation of the KKT necessary conditions 5.3. Then we have

$$\frac{d\theta(\mu)}{d\mu} = -\sum_{i=1}^{n} \ln x_i. \tag{5.9}$$

Finally, we can write the rate of change of the barrier parameter μ by using equation 5.6 as follows

$$\frac{d\mu}{dt} = \sum_{i=1}^{n} \ln x_i. \tag{5.10}$$

Now, by using equations 5.10 and 5.3, we can determine the central trajectory. The differential equation 5.10 and algebraic equation system 5.3 form a system of Differential-Algebraic Equations (DAEs) for problem P. The DAE can be written as follows:

$$\begin{aligned} \frac{d\mu}{dt} &= \sum_{i=1}^{n} \ln x_i \\ Ax - b &= 0 \\ A^T y + z &= c \\ XZe - \mu e &= 0. \end{aligned} \tag{5.11}$$

To find a solution to the DAE, we need to state the following theorem.

Theorem 5.2 *The DAE defined by 5.11 has index 1.*

Proof 5.2 Let us rewrite the DAE in semi-explicit form as in equation 4.3

$$\begin{aligned} \frac{d\mu}{dt} &= f(x, y, z, \mu, t) \\ 0 &= g(x, y, z, \mu, t), \end{aligned} \tag{5.12}$$

where $g(x, y, z, \mu, t) = \begin{pmatrix} Ax - b \\ A^T y + z - c \\ ZXe - \mu e \end{pmatrix}$. Then by differentiating the algebraic equation of 5.12 with respect to t, we get

$$g_x \frac{dx}{dt} + g_y \frac{dy}{dt} + g_z \frac{dz}{dt} + g_\mu \frac{d\mu}{dt} = 0, \tag{5.13}$$

where $g_x = \begin{bmatrix} A \\ 0 \\ Z \end{bmatrix}$, $g_y = \begin{bmatrix} 0 \\ A^T \\ 0 \end{bmatrix}$, $g_z = \begin{bmatrix} 0 \\ I \\ X \end{bmatrix}$, $g_\mu = \begin{bmatrix} 0 \\ 0 \\ -e \end{bmatrix}$.

With some algebraic manipulation, equation 5.13 becomes

$$\begin{bmatrix} g_x & g_y & g_z \end{bmatrix} \begin{bmatrix} \frac{dx}{dt} & \frac{dy}{dt} & \frac{dz}{dt} \end{bmatrix}^T = -g_\mu \frac{d\mu}{dt}. \tag{5.14}$$

The RHS of 5.14 becomes $h(x, y, z, \mu, t)$ because of equation 5.10. Note that, $[g_x, g_y, g_z]$ is equal to $J(x, y, z)$ as defined in equation 3.12, that is the Jacobian of the nonlinear system 3.9. Next, we need coordinate changes.

Let $v = (x, y, z)$ and $\frac{dv}{dt} = (\frac{dx}{dt}\ \frac{dy}{dt}\ \frac{dz}{dt})$. Then equation 5.14 becomes

$$J(v)\frac{dv}{dt} = f(v, \mu, t), \tag{5.15}$$

where $f(v, \mu, t) = -g_\mu \frac{d\mu}{dt}$. By the assumption in the previous section, the Jacobian $J(v)$ is nonsingular. That guarantees that $J(v)$ is invertible. Finally, we will have the following implicit ODE

$$\frac{dv}{dt} = J^{-1}(v) f(v, \mu, t) = F(v, \mu, t). \tag{5.16}$$

By the definition of index as it is given in [12], an implicit ODE resulted after one differentiation step. This implies that the DAE defined by 5.11 has index 1. This concludes the proof.

Semi-explicit DAEs of index 1 can be solved by using the techniques for ODEs. The algebraic equations in the DAE can be further transformed into differential equations. Differentiating the algebraic equations 5.3 with respect to t, we get

$$\begin{aligned} A\frac{dx}{dt} &= 0 \\ A^T\frac{dy}{dt} + \frac{dz}{dt} &= 0 \\ X\frac{dz}{dt} + Z\frac{dx}{dt} &= \frac{d\mu}{dt}e. \end{aligned} \tag{5.17}$$

With some algebraic manipulation by using the sparsity of system, the solution to equation 5.15 is as follows

$$\begin{aligned} \frac{dx}{dt} &= \frac{d\mu}{dt}\left[I - Z^{-1}XA^T(AZ^{-1}XA^T)^{-1}A\right]Z^{-1}e \\ \frac{dy}{dt} &= -\frac{d\mu}{dt}(AZ^{-1}XA^T)^{-1}AZ^{-1}e \\ \frac{dz}{dt} &= \frac{d\mu}{dt}A^T(AZ^{-1}XA^T)^{-1}AZ^{-1}e. \end{aligned} \tag{5.18}$$

By taking the initial values for x, y, and z which satisfy the algebraic equations of the DAE in the interior of the feasible region, and taking a small initial value for μ, we can solve 5.10 and 5.18.

5.4 A Modified Primal Dual IPM for LO Problems

In this section, we present a primal dual IPM for LO problems that solves the DAE in equation 5.11 to determine the central trajectory of the LO problem. By following the central trajectory, we have found an optimal solution to the LO problem. Let us consider the same problem P as defined in equation 5.1. To solve the DAE system in equation 5.11, we use the backward Euler's method for simplicity. In Euler's method, the ODE that is defined in equation 5.10 can be approximated as follows

$$\mu^k - \mu^{k-1} = h\sum_{i=1}^{n} \ln {x_i}^k, \tag{5.19}$$

where h is a given step length. By applying Euler's method for the DAE system in equation 5.11, we obtain the following nonlinear system

$$\begin{aligned}
\mu^k - \mu^{k-1} - h\sum_{i=1}^{n} \ln {x_i}^k &= 0 \\
Ax^k - b &= 0 \\
A^T y^k + z^k - c &= 0 \\
X^k Z^k e - \mu^k e &= 0.
\end{aligned} \tag{5.20}$$

Newton's method is used to find a solution to this problem. Given $x^0 > 0, y^0$ and $z^0 > 0$ and $\mu^0 > 0$, moving direction vectors dx, dy, dz, and $d\mu$ that move from the current point to a new point while satisfying the DAE system in equation 5.11, are determined by using the following linearized system

$$\begin{bmatrix} 0 & 0 & 0 & e \\ Z & 0 & X & -e \\ A & 0 & 0 & 0 \\ 0 & A^T & I & 0 \end{bmatrix} \begin{bmatrix} dx \\ dy \\ dz \\ d\mu \end{bmatrix} = - \begin{bmatrix} -h\sum_{i=1}^{n} \ln {x_i}^0 e \\ X^0 Z^0 e - \mu^0 e \\ Ax^0 - b \\ A^T y^0 + z^0 - c \end{bmatrix}. \tag{5.21}$$

Let us define $d_P = Ax^0 - b$ and $d_D = A^T y^0 + z^0 - c$. Then the solution of the linear system 5.21 with respect to the moving direction vectors dx, dy, dz, and $d\mu$ can be found as follows. By addition of the first two equations we have

$$Zdx + Xdz = \mu^0 e - X^0 Z^0 e + h\sum_{i=1}^{n} \ln {x_i}^0 e. \tag{5.22}$$

Let us define $d = \mu^0 e - X^0 Z^0 e + h \sum_{i=1}^{n} \ln {x_i}^0 e$. From here dx can be expressed as follows

$$dx = Z^{-1}(d - Xdz). \tag{5.23}$$

From the last equation, we have

$$dz = d_D - A^T dy. \tag{5.24}$$

Finally, by using the third equation of 5.21, and equations 5.23 and 5.24 we have

$$dy = (AZ^{-1}XA^T)^{-1}(AZ^{-1}Xd_D + d_P - AZ^{-1}d), \tag{5.25}$$

and

$$d\mu = h \sum_{i=1}^{n} \ln {x_i}^0. \tag{5.26}$$

So far, we have computed the moving direction vectors dx, dy, dz and $d\mu$ by using the current points (x^0, y^0, z^0), and μ^0. The new points for the next iteration can be obtained as follows:

$$\begin{aligned} x^1 &= x^0 + \rho\alpha_P dx \\ y^1 &= y^0 + \rho\alpha_D dy \\ z^1 &= z^0 + \rho\alpha_D dz \\ \mu^1 &= \mu^0 + d\mu, \end{aligned} \tag{5.27}$$

where $0 < \rho < 1$, and α_P and α_D are step sizes for the primal and dual problem respectively. Step sizes are chosen to assure that x and z are positive or feasible as follows

$$\alpha_P = \min\left\{-\frac{x_i}{dx_i} : \forall\ dx_i < 0, \quad 1 \le i \le n\right\}$$

$$\alpha_D = \min\left\{-\frac{z_i}{dz_i} : \forall \ dz_i < 0, \quad 1 \le i \le n\right\}. \tag{5.28}$$

From given initial solution points (x^0, y^0, z^0), and μ^0, we have computed the new points (x^1, y^1, z^1) and μ^1 such that $\mu^1 < \mu^0$ by solving the DAE system in 5.11. The proposed primal dual IPM continues to iterate by reducing μ as explained in equation 5.19 and computing moving direction vectors dx, dy, dz and $d\mu$ and the new points until convergence is achieved. Convergence of the proposed primal dual IPM can be achieved like in the traditional primal dual IPMS by making the duality gap to be close to zero with a given accuracy ε. Specifically we need to satisfy

$$c^T x - b^T y = \varepsilon. \tag{5.29}$$

The modified primal dual IPM is similar to the classical primal dual IPMs. The differences are on the barrier parameter update procedure and the moving directions. Note that, we are using similar formulas as in the classical primal dual IPM, but the directions are different. The reason is that, we have a different parameterization of the barrier parameter μ than in the traditional primal dual IPM that was explained in Chapter 3. These differences affect the duality gap that is the stopping criteria of the primal dual IPM. As a result, we can state that we have a different duality gap reduction than the classical approaches. A modified primal dual IPM can be stated as follows.

Proposed Primal Dual IPM Algorithm for LO Problems

Step 0 Initialization:

- SET iteration number $k = 0, \varepsilon, \rho$ and h.
- START with an initial solution $x^k > 0, y^k, z^k > 0$ and $\mu^k > 0$.

Step 1 Stopping Criteria:

- IF equation 5.29 is satisfied GO TO Step 3.
- ELSE GO TO Step 2.

Step 2 Iteration:

- COMPUTE moving directions $dx^k, dy^k, dz^k, d\mu^k$ by using equations 5.23, 5.25, 5.24, and 5.26 respectively.
- COMPUTE step sizes α_P and α_D by using equation 5.28
- MOVE to the next solution $x^{k+1}, y^{k+1}, z^{k+1}$ using equation 5.27
- UPDATE the barrier parameter μ^k using equation 5.19
- SET $k = k + 1$
- GO TO Step 1.

Step 3 Stop: with the optimal point $x^*, y^*, z^* = x^k, y^k, z^k$.

5.5 New Parameterization of the Central Trajectory for Convex QO Problems

Let us consider the primal QO problem QP in the standard form

$$QP: \quad \begin{array}{rcl} \min \quad f(x) & = & c^T x + \frac{1}{2} x^T Q x \\ \text{subject to} & & \\ Ax & = & b \\ x & \geq & 0, \end{array} \tag{5.30}$$

where $x \in \Re^n$ is a vector of decision variables, $A \in \Re^{m \times n}$ is a matrix of constraints, $Q \in \Re^{n \times n}$ is a symmetric positive semi definite coefficient matrix of the quadratic terms and $c \in \Re^n$ is a coefficient vector of the linear terms of the objective function $f(x)$, and $b \in \Re^m$ is a RHS vector of constraints.

Adding a barrier term to the objective function of QP results to the following problem QP_μ.

$$QP_\mu: \quad \begin{array}{rcl} \min \quad f(x, \mu) & = & c^T x + \frac{1}{2} x^T Q x - \mu \sum_{i=1}^{n} \ln x_i \\ \text{subject to} & & \\ Ax & = & b \\ x & > & 0, \end{array} \tag{5.31}$$

where $\mu > 0$ is the barrier parameter.

The associated KKT necessary conditions give the following system

$$Ax - b = 0$$

$$\begin{aligned} -Qx + A^T y + z - c &= 0 \\ XZe - \mu e &= 0 \end{aligned} \tag{5.32}$$

The problem QP_μ can also be defined as follows:

$$\theta(\mu) = \inf \left\{ c^T x + \frac{1}{2} x^T Q x - \mu \sum_{i=1}^{n} \ln x_i \quad \text{such that} \quad Ax = b, x > 0 \right\} \tag{5.33}$$

By a similar approach to the LO case, from the theorem 5.1 by Bazaraa et al. [7] (Theorem 9.4.3), it is stated that the optimal solution to the problem QP could be obtained by minimizing $\theta(\mu)$. It should be noted that the quadratic problem is strictly convex and hence this theorem holds. In other words,

$$\text{minimum} \left\{ c^T x + \frac{1}{2} x^T Q x \quad \text{such that} \quad Ax = b, x \geq 0 \right\} = \inf_{\mu > 0} \theta(\mu). \tag{5.34}$$

Note that, 5.30 and 5.33 are strictly convex optimization problems because both the objective function is strictly convex and the constraints are convex. Consequently, for any fixed $\mu, \theta(\mu)$ has a unique minimum. The minimum of $\theta(\mu)$ will be found by using the steepest descent method. Again, we consider μ as a function of a parameter t. To find the rate of change of the barrier parameter μ, we have to move in the direction of the negative gradient of $\theta(\mu)$. Thus,

$$\frac{d\mu}{dt} = -\frac{d\theta(\mu)}{d\mu}. \tag{5.35}$$

From equation 5.33, $\theta(\mu)$ is differentiable. Now, let us find its derivative with respect to μ.

$$\frac{d\theta(\mu)}{d\mu} = c^T \frac{dx}{d\mu} + x^T Q \frac{dx}{d\mu} - \sum_{i=1}^{n} \ln x_i - \mu \sum_{i=1}^{n} \frac{1}{x_i} \frac{dx_i}{d\mu} \quad \text{such that} \quad A\frac{dx}{d\mu} = 0 \tag{5.36}$$

If we consider the second equation from equation 5.32 $c^T = y^T A + z^T - x^T Q$ and substitute in 5.36 then we have

$$\frac{d\theta(\mu)}{d\mu} = z^T \frac{dx}{d\mu} - \sum_{i=1}^{n} \ln x_i - \mu \sum_{i=1}^{n} \frac{1}{x_i} \frac{dx_i}{d\mu} \quad \text{such that} \quad A\frac{dx}{d\mu} = 0. \tag{5.37}$$

We can see that the first and the last terms cancel each other by the help of the third equation of KKT necessary conditions. Then we have

$$\frac{d\theta(\mu)}{d\mu} = -\sum_{i=1}^{n} \ln x_i. \tag{5.38}$$

Finally, we can write the rate of change of the barrier parameter by using equation 5.35 μ as follows.

$$\frac{d\mu}{dt} = \sum_{i=1}^{n} \ln x_i \tag{5.39}$$

Now, by using equations 5.39 and 5.32, we can determine the central trajectory. The differential equation 5.39 and the algebraic equation system 5.32 form a system of DAEs for the problem QP. The DAE can be written as follows:

$$\begin{aligned} \frac{d\mu}{dt} &= \sum_{i=1}^{n} \ln x_i \\ Ax - b &= 0 \\ -Qx + A^T y + z - c &= 0 \\ XZe - \mu e &= 0. \end{aligned} \tag{5.40}$$

To find a solution to the DAE, we need to state the following theorem.

Theorem 5.3 *The DAE defined by 5.40 has index 1.*

Proof 5.3 It is similar to the proof of theorem 5.2.

The algebraic equations in the DAE can be further transformed into differential equations. Differentiating the algebraic equations 5.40 we get

$$\begin{aligned} A\frac{dx}{dt} &= 0 \\ -Q\frac{dx}{dt} + A^T\frac{dy}{dt} + \frac{dz}{dt} &= 0 \\ X\frac{dz}{dt} + Z\frac{dx}{dt} &= \frac{d\mu}{dt}e. \end{aligned} \tag{5.41}$$

With some algebraic manipulation, the solution to this system is as follows

$$\begin{aligned} \frac{dx}{dt} &= \frac{d\mu}{dt}\left[I - EXA^T(AEXA^T)^{-1}A\right]Ee \\ \frac{dy}{dt} &= -\frac{d\mu}{dt}(AEXA^T)^{-1}AEe \\ \frac{dz}{dt} &= \frac{d\mu}{dt}\left\{\left[Q - QEXA^T(AEXA^T)^{-1}A\right] + A^T(AEXA^T)^{-1}A\right\}Ee, \\ \text{where} \quad E &= (Z + XQ)^{-1}. \end{aligned} \tag{5.42}$$

By taking the initial values for x, y, and z which satisfy the algebraic equations of the DAE, being in the interior of the feasible region, and taking a small initial value for μ, we can solve 5.39 and 5.42.

5.6 A Modified Primal Dual IPM for Convex QO Problems

In this section, we present a primal dual IPM for Convex QO problems that solves the DAE in equation 5.40 to determine the central trajectory of the LO problem. By following the central trajectory, we have found an optimal solution to the LO problem. Let us consider the same problem QP as defined in equation 5.30. To solve the DAE system in equation 5.40, we use the backward Euler's method for simplicity. In Euler's method, the ODE that is defined in equation 5.10 can be approximated as follows

$$\mu^k - \mu^{k-1} = h \sum_{i=1}^{n} \ln x_i, \tag{5.43}$$

where h is a given step length. By applying Euler's method for the DAE system in equation 5.40, we obtain the following nonlinear system

$$\begin{aligned} \mu^k - \mu^{k-1} - h \sum_{i=1}^{n} \ln {x_i}^k &= 0 \\ Ax^k - b &= 0 \\ -Qx^k + A^T y^k + z^k - c &= 0 \\ X^k Z^k e - \mu^k e &= 0. \end{aligned} \tag{5.44}$$

Again, Newton's method is used to find a solution to this nonlinear system. Given $x^0 > 0, y^0$ and $z^0 > 0$ and $\mu^0 > 0$, moving direction vectors dx, dy, dz, and $d\mu$ those move from the current point to a new point while satisfying the DAE system in equation 5.40, are determined by solving the following linearized system in a similar way to LO case

$$\begin{bmatrix} 0 & 0 & 0 & e \\ Z & 0 & X & -e \\ A & 0 & 0 & 0 \\ -Q & A^T & I & 0 \end{bmatrix} \begin{bmatrix} dx \\ dy \\ dz \\ d\mu \end{bmatrix} = - \begin{bmatrix} -h \sum_{i=1}^{n} \ln {x_i}^0 e \\ X^0 Z^0 e - \mu^0 e \\ Ax^0 - b \\ -Qx^0 + A^T y^0 + z^0 - c \end{bmatrix}. \tag{5.45}$$

Let us define $d_P = Ax^0 - b$ and $d_D = -Qx^0 + A^T y^0 + z^0 - c$. Then the solution of the linear system 5.45 with respect to the moving direction vectors dx, dy, dz, and $d\mu$

can be found as before for LO case. By adding the first two equations, we have

$$Zdx + Xdz = \mu^0 e - X^0 Z^0 e + h \sum_{i=1}^{n} \ln x_i{}^0 e. \tag{5.46}$$

Let us define $d = \mu^0 e - X^0 Z^0 e + h \sum_{i=1}^{n} \ln x_i{}^0 e$. Then equation 5.46 can be written as follows

$$Zdx + Xdz = d. \tag{5.47}$$

By multiplying the last equation of 5.45 by X, we have

$$-XQdx + XA^T dy + Xdz = Xd_D. \tag{5.48}$$

Subtracting equation 5.47 from equation 5.48 would result

$$-(Z + XQ)dx + XA^T dy = Xd_D - d. \tag{5.49}$$

Multiplying equation 5.49 by $A(Z + XQ)^{-1}$, we get

$$-Adx + A(Z + XQ)^{-1} XA^T dy = A(Z + XQ)^{-1}(Xd_D - d). \tag{5.50}$$

Adding the third equation of 5.45 to equation 5.50, we have

$$A(Z + XQ)^{-1} XA^T dy = A(Z + XQ)^{-1}(Xd_D - d) + d_P. \tag{5.51}$$

From here dy can be expressed as follows:

$$dy = [A(Z + XQ)^{-1} XA^T]^{-1}[A(Z + XQ)^{-1}(Xd_D - d) + d_P]. \tag{5.52}$$

By equation 5.49 we further have

$$dx = (Z + XQ)^{-1}[X(A^T dy - d_D) + d]. \tag{5.53}$$

Also from equation 5.47 we get

$$dz = X^{-1}(d - Zdx), \tag{5.54}$$

and

$$d\mu = h \sum_{i=1}^{n} \ln {x_i}^0. \tag{5.55}$$

So far, we have computed the moving direction vectors dx, dy, dz and $d\mu$ by using the current points (x^0, y^0, z^0), and μ^0. The new points for the next iteration can be obtained as follows:

$$\begin{aligned} x^1 &= x^0 + \rho\alpha_P dx \\ y^1 &= y^0 + \rho\alpha_D dy \\ z^1 &= z^0 + \rho\alpha_D dz \\ \mu^1 &= \mu^0 + d\mu, \end{aligned} \tag{5.56}$$

where $0 < \rho < 1$, and α_P and α_D are step sizes for the primal and dual problem respectively. Step sizes are chosen to assure that x and z are positive or feasible as follows

$$\begin{aligned} \alpha_P &= \min\left\{-\frac{x_i}{dx_i} : \forall \ dx_i < 0, \quad 1 \le i \le n\right\} \\ \alpha_D &= \min\left\{-\frac{z_i}{dz_i} : \forall \ dz_i < 0, \quad 1 \le i \le n\right\}. \end{aligned} \tag{5.57}$$

From given initial solution points (x^0, y^0, z^0), and μ^0, we have computed the new points (x^1, y^1, z^1) and μ^1 such that $\mu^1 < \mu^0$ by solving the DAE system in 5.40. The proposed primal dual IPM continues to iterate by reducing μ as explained in equation 5.43 and computing moving direction vectors dx, dy, dz and $d\mu$ and the new points until convergence is achieved. Convergence of the the proposed primal dual IPM can be achieved like in the traditional primal dual IPMS by making the duality gap to be close to zero with a given accuracy ε. Specifically, we need to satisfy

$$c^T x - b^T y = \varepsilon. \tag{5.58}$$

The modified primal dual IPM is similar to the classical primal dual IPMs. The differences are on the barrier parameter update procedure and the moving directions. Note that, we are using similar formulas as in the classical primal dual IPM, but the directions are different. The reason is that, we have a different parameterization of the barrier parameter μ than in the classical primal dual IPM that was explained in Chapter 3. These differences affect the duality gap that is the stopping criteria of the primal dual IPM. As a result, we can state that we have a different duality gap reduction than the classical approaches. A modified primal dual IPM can be stated as follows.

Proposed Primal Dual IPM Algorithm for QO Problems

Step 0 Initialization:

- SET iteration number $k = 0, \varepsilon, \rho$ and h.
- START with an initial solution $x^k > 0, y^k, z^k > 0$ and $\mu^k > 0$.

Step 1 Stopping Criteria:

- IF equation 5.58 is satisfied GO TO Step 3.
- ELSE GO TO Step 2.

Step 2 Iteration:

- COMPUTE moving directions $dx^k, dy^k, dz^k, d\mu^k$ by using equations 5.53, 5.52, 5.54, and 5.55 respectively.
- COMPUTE step sizes α_P and α_D using equation 5.57
- MOVE to the next solution $x^{k+1}, y^{k+1}, z^{k+1}$ using equation 5.56
- UPDATE the barrier parameter μ^k using equation 5.47
- SET $k = k + 1$
- GO TO Step 1.

Step 3 Stop: with the optimal point $x^*, y^*, z^* = x^k, y^k, z^k$.

Chapter 6

Convergence, Implementation, Computational Complexity and Stability Details

6.1 Convergence Details

In this section, we study the convergence of the proposed primal dual IPMs by using the new parameterization of the central trajectory. Let us consider the following primal LO problem P in the standard form and the corresponding dual problem D.

$$\begin{aligned} P: \quad & \min \quad f(x) = c^T x \\ & \text{subject to} \\ & \qquad Ax = b \end{aligned} \tag{6.1}$$

$$x \geq 0,$$

where $x \in \Re^n$, $A \in \Re^{m\times n}$, $c \in \Re^n$, and $b \in \Re^m$.

The dual problem D of problem P becomes

$$\begin{aligned} D: \qquad \max \quad g(y) &= b^T y \\ \text{subject to} & \qquad\qquad (6.2) \\ A^T y + z &= c \\ z &\geq 0, \end{aligned}$$

where $y \in \Re^m$ and $z \in \Re^n$.

Our first objective is to show that the sequence of points $\{x^k, y^k$, and $z^k\}$ generated by the proposed primal dual IPM using the new parameterization of the central trajectory converges to optimal solution of the problem. Then, second objective is to show that the convergence of the new parameterization is faster than original parameterization. Before we study the convergence, the following assumptions for problem P and its dual problem D are necessary.

1. The feasible region S of problem P is not empty. Therefore there exists an $x \in S$.
2. The feasible region T of problem D is not empty. Therefore there exists an $(y, z) \in T$.
3. A has a full row-rank. $Rank(A) = m \leq n$

Now, we begin to study the convergence of the proposed primal dual IPMs by using the new parameterization of the central trajectory by explaining some properties of the trajectories of the solution. We start with the basic facts.

First, we need to show that $x \in S$ and $(y, z) \in T$ determined by the first order KKT conditions are continuous functions of μ and t. The corresponding KKT conditions can be written as follows:

$$\begin{aligned} F_1(x, y, z, \mu) = Ax - b &= 0 \\ F_2(x, y, z, \mu) = A^T y + z - c &= 0 \\ F_3(x, y, z, \mu) = XZe - \mu e &= 0 \end{aligned}$$

and

$$F(x, y, z, \mu) = [F_1(x, y, z, \mu), F_2(x, y, z, \mu), F_3(x, y, z, \mu)]^T = 0. \tag{6.3}$$

In order to show that x, y, and z are continuous functions of μ, the following implicit function theorem [36] is to be satisfied.

Theorem 6.1 Implicit Function Theorem: *For $i = 1, \ldots, m$, let the functions $F_i(y_1, \ldots, y_m, x_1, \ldots, x_n)$ all be defined in a neighborhood of the point P_k : $({y_1}^k, \ldots, {y_m}^k, {x_1}^k, \ldots, {x_n}^k)$ and have continuous first partial derivatives in this neighborhood. For $i = 1, \ldots, m$, let the equations $F_i(y_1, \ldots, y_m, x_1, \ldots, x_n) = 0$ be satisfied at P_k and let*

$$\frac{\partial(F_1, \ldots, F_m)}{\partial(y_1, \ldots, y_m)} \neq 0 \text{ at } P_k.$$

Then, in an appropriate neighborhood of $({x_1}^k, \ldots, {x_n}^k)$ there is a unique set of continuous functions $y_i = f_i(x_1, \ldots, x_n)$ for $i = 1, \ldots, m$, such that ${y_i}^k = f_i({x_1}^k, \ldots, {x_n}^k)$

for $i = 1, \ldots, m$, *and for all* i, $F_i(f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n), x_1, \ldots, x_n) = 0$ *in that neighborhood.*

In order to apply the implicit function theorem, the determinant of the Jacobian $J(x, y, z)$ of the nonlinear system in 6.3 at the given $x^k \in S$, $y^k, z^k \in T$ is to be calculated. Specifically,

$$\left|J(x^k, y^k, z^k)\right| = \begin{vmatrix} Z_k & 0 & X_k \\ A & 0 & 0 \\ 0 & A^T & I \end{vmatrix} = \begin{vmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{vmatrix}$$

$$= |J_{11}| \left|J_{22} - J_{21} J_{11}^{-1} J_{12}\right|,$$

where

$$J_{11} = Z_k, J_{12} = [0, X_k], J_{21} = \begin{bmatrix} A \\ 0 \end{bmatrix}, J_{22} = \begin{bmatrix} 0 & 0 \\ A^T & I \end{bmatrix}.$$

Clearly,

$$J_{21} J_{11}^{-1} J_{12} = \begin{bmatrix} 0 & A{Z_k}^{-1} X_k \\ 0 & 0 \end{bmatrix}$$

and

$$J_{22} - J_{21} J_{11}^{-1} J_{12} = \begin{bmatrix} 0 & -A{Z_k}^{-1} X_k \\ A^T & I \end{bmatrix}.$$

Thus

$$\left|J_{22} - J_{21} J_{11}^{-1} J_{12}\right| = \left|A{Z_k}^{-1} X_k A^T\right|$$

and

$$\left|J(x^k, y^k, z^k)\right| = \left|Z^k\right| \left|AZ_k{}^{-1}X_kA^T\right| = \frac{1}{\mu_k}\left|Z_k\right| \left|X_0{}^2\right| \left|AA^T\right| \neq 0.$$

Note that, A has full row rank by the assumption, hence $AA^T \neq 0$ and $x^k \in S$, $y^k, z^k \in T$ implies that $|X_k| > 0$ and $|Z_k| > 0$. By the implicit function theorem, since $\left|J(x^k, y^k, z^k)\right| \neq 0$, there are unique set of continuous functions $x(\mu)$, $y(\mu)$ and $z(\mu)$ passing through $x^k = x(\mu^k)$, $y^k = y(\mu^k)$ and $z^k = z(\mu^k)$.

After showing x, y, and z are continuous functions of μ, next we need to show that x, y, and z are also continuous functions of t. In the previous chapter we defined the rate of change for the barrier parameter μ as

$$\frac{d\mu}{dt} = \sum_{i=1}^{n} \ln x_i.$$

It is obvious that $\sum \ln x_i$ is a continuous function of x and, x is continuous function of μ. Therefore we conclude that the rate of change for the barrier parameter μ is also continuous in t by using the theorem of existence and uniqueness of solutions of nonlinear differentiable equations [36]. Therefore x, y and z are also continuous function of t. That concludes the work on continuity.

Next, we need to state the following convergence theorem for the proposed primal dual IPMs by using the new parameterization of the central trajectory and prove it.

Theorem 6.2 Convergence Theorem *Let $x^0 \in S$ and $(y^0, z^0) \in T$ and μ^0 are given. Suppose that $\bar{x}(t), \bar{y}(t), \bar{z}(t)$, and $\mu(t)$ denote the trajectories of the solution of the system of equations in 5.10 and 5.18 with given initial solution (x^0, y^0, z^0, μ^0). Then, either $\bar{x}(t) \in S$, $(\bar{y}(t), \bar{z}(t)) \in T$ and $\mu(t) \geq 0$, or $\bar{x}(t), \bar{y}(t)$, and $\bar{z}(t)$ correspond* *to the optimal solution to the given problem.*

Proof 6.2 We need to show that with the proposed parameterization of the central trajectory of the solution is either at the optimal solution or in the feasible region. Let us consider the DAE system in 5.11. Since $\mu_0 > 0$ and $\mu(t)$ is continuous in t, before $\mu(t)$ reaches a negative value, it must be $\mu(t) = 0$ at some time t, which implies that complementary slackness conditions of KKT conditions are satisfied. In other words,

$$\bar{x}_j(t)\bar{z}_j(t) = 0, j = 1, \ldots, n.$$

Then, we need to show that $\bar{x}(t), \bar{y}(t)$, and $\bar{z}(t)$ satisfy the primal and dual feasibility conditions of the KKT conditions. In order to do this, we need to show the solutions of the system of the differential equations 5.17 satisfy the KKT conditions

$$\begin{aligned} A\frac{d\bar{x}}{dt} &= 0 \\ A^T\frac{d\bar{y}}{dt} + \frac{d\bar{z}}{dt} &= 0 \\ X\frac{d\bar{z}}{dt} + Z\frac{d\bar{x}}{dt} &= \frac{d\mu}{dt}e. \end{aligned} \tag{6.4}$$

Integrating both sides of the first equation, we obtain

$$A\int_0^t \frac{d\bar{x}}{d\sigma}d\sigma = 0.$$

It follows from here that

$$A\bar{x}(t) - A\bar{x}(0) = 0.$$

But

$$A\bar{x}(0) = b.$$

Thus,

$$A\bar{x}(t) = b.$$

We can state that primal feasibility is satisfied. Similarly, we can show that dual feasibility is also satisfied as follows

$$A^T \int_0^t \frac{d\bar{y}}{d\sigma} d\sigma + \int_0^t \frac{d\bar{z}}{d\sigma} d\sigma = 0.$$

It follows from here that

$$A^T \bar{y}(t) - A^T \bar{y}(0) + \bar{z}(t) - \bar{z}(0) = 0,$$

but

$$A^T \bar{y}(0) + \bar{z}(0) = c.$$

Thus,

$$A^T \bar{y}(t) + \bar{z}(t) = c.$$

That concludes that dual feasibility satisfied.

We see that $\bar{x}(t), \bar{y}(t)$, and $\bar{z}(t)$ satisfy the KKT conditions. Therefore the optimality conditions for the optimization problem are satisfied. This shows that $\bar{x}(t), \bar{y}(t)$, and $\bar{z}(t)$ correspond to the optimal solution.

Next we need to show that the trajectory is in the feasible region. In other words, $\bar{x}(t) \in S$ and $(\bar{y}(t), \bar{z}(t)) \in T$. Since $\bar{x}_j(t)\bar{z}_j(t) = \mu(t)$, and $\mu(t) > 0$ implies that either $\bar{x}_i(t) > 0$ and $\bar{z}_i(t) > 0$ or $\bar{x}_i(t) < 0$ and $\bar{z}_i(t) < 0$. As one can see they can not be negative. Otherwise, because $\bar{x}_i(t) > 0$ and $\bar{z}_i(t) > 0$, $\bar{x}_i(t)$, and $\bar{z}_i(t)$ are continuous in t, before $\bar{x}_i(t)$, and $\bar{z}_i(t)$ become negative there exist a time t^* such that $\bar{x}_j(t^*)\bar{z}_j(t^*) = 0$, that implies that we have optimal solution to the given problem P.

Now we consider $\bar{x}_i(t) > 0$ and $\bar{z}_i(t) > 0$. Previously, we showed that $A\bar{x}(t) = b$ and $A^T\bar{y}(t) + \bar{z}(t) = c$. Thus, $\bar{x}(t) \in S$ and $(\bar{y}(t), \bar{z}(t)) \in T$. This means that the trajectory is in the sets S and T and in the interior of the feasible region. This completes the proof.

We can always compute the duality gap and the duality gap is reduced at each iteration. As we said before, convergence of the proposed primal dual IPMs by using the new parameterization of the central trajectory has been achieved when the duality gap becomes zero or close to zero with the given accuracy. Next, we show that the proposed primal dual IPMs by using the new parameterization of the central trajectory reduces the duality gap at each iteration point and it converges to the optimal solution when the duality gap becomes zero or close to zero. In addition to this, we show that the convergence rate for the proposed primal dual IPMs by using the new parameterization of the central trajectory is faster than classical primal dual IPMs. Now, let us start to investigate the duality gap for the classical primal dual IPMs.

Let assume that at kth iteration of the algorithm we have primal feasibility, $Ax^k = b$ and dual feasibility, $A^Ty^k + z^k = c$. Let $gap(k)$ corresponds to the current duality gap at the kth iteration.

$$\begin{aligned} gap(k) &= c^T x^k - b^T y^k \\ &= (A^T y^k + z^k)^T x^k - (Ax^k)^T y^k \\ &= {z^k}^T x^k \end{aligned}$$

Let α be the minimum of the primal and dual step sizes, $\alpha = min(\alpha_P, \alpha_D)$. Then as we move

$$\begin{aligned} gap(k+1) &= (z^k + \alpha dz)^T (x^k + \alpha dx) \\ &= {z^k}^T x^k + \alpha({z^k}^T dx + dz^T x^k) + \alpha^2 dz^T dx. \end{aligned}$$

As we remember, primal and dual feasibility imply $dz = -A^T dy$ and $Adx = 0$ so $dz^T dx = 0$. Thus

$$\begin{aligned} gap(k+1) &= {z^k}^T x^k + \alpha \left[(Z^k dx)^T e + (X^k dz)^T e \right] \\ &= gap(k) + \alpha(Z^k dx + X^k dz)^T e. \end{aligned}$$

Then by using the linearized complementary slackness or $Zdx + Xdz = \mu e - X^k Z^k e$ from equation 3.13, we get

$$\begin{aligned} gap(k+1) &= gap(k) + \alpha(\mu e - X^k Z^k e)^T e \\ &= gap(k) + \alpha(n\mu - gap(k)). \end{aligned}$$

We see that the duality gap is reduced if $gap(k+1) < gap(k)$ as long as

$$\mu < \frac{gap(k)}{n}. \tag{6.5}$$

For the classical primal dual IPMs, the most commonly used update rule for the barrier parameter μ is

$$\mu^{k+1} < \mu^k(1-\sigma), 0 < \sigma < 1.$$

This is a discrete update rule and it can be extracted to the following continuous form.

$$\frac{d\mu}{dt} = -\sigma\mu$$

Thus,

$$\mu(t) = e^{-\sigma\mu}. \tag{6.6}$$

Next, we investigate the duality gap for the proposed primal dual IPMs by using the new parameterization of the central trajectory.

Let us assume that at the kth iteration of the proposed algorithm, we have primal feasibility, $Ax^k = b$ and dual feasibility, $A^Ty^k + z^k = c$. Let $gap(k)$ corresponds to the current duality gap at the kth iteration.

$$\begin{aligned} gap(k) &= c^Tx^k - b^Ty^k \\ &= (A^Ty^k + z^k)^Tx^k - (Ax^k)^Ty^k \\ &= {z^k}^Tx^k \end{aligned}$$

Let α be the minimum of primal and dual step sizes, $\alpha = min(\alpha_P, \alpha_D)$. Then as we move along the central trajectory

$$\begin{aligned} gap(k+1) &= (z^k + \alpha dz)^T(x^k + \alpha dx) \\ &= {z^k}^Tx^k + \alpha({z^k}^Tdx + dz^Tx^k) + \alpha^2 dz^Tdx. \end{aligned}$$

As we remember, primal and dual feasibility imply $dz = -A^Tdy$ and $Adx = 0$. Therefore $dz^Tdx = 0$. Thus

$$\begin{aligned} gap(k+1) &= {z^k}^Tx^k + \alpha\left[(Z^kdx)^Te + (X^kdz)^Te\right] \\ &= gap(k) + \alpha(Z^kdx + X^kdz)^Te. \end{aligned}$$

Then by using the linearized complementary slackness or $Zdx + Xdz = \mu e - X^k Z^k e + h\sum_{i=1}^{n} \ln x_i{}^0 e$ from equation 5.22, we get

$$\begin{aligned} gap(k+1) &= gap(k) + \alpha(\mu e - X^k Z^k e + h\sum_{i=1}^{n} \ln x_i{}^0 e)^T e \\ &= gap(k) + \alpha(n\mu + h\sum_{i=1}^{n} \ln x_i{}^0 e - gap(k)). \end{aligned}$$

As we see that the duality gap is reduced if $gap(k+1) < gap(k)$ as long as

$$\mu < \frac{gap(k) - h\sum_{i=1}^{n} \ln x_i{}^0}{n}. \tag{6.7}$$

We see that we have an additional term in 6.7. This term puts a bound to the gap reduction. If we compare 6.7 with 6.5, we can conclude that for the proposed primal dual IPMs by using the new parameterization of the central trajectory, the bound to the gap reduction results into faster convergence to 0 for μ and the duality gap. We know that, when $\mu \longrightarrow 0$ the duality gap converges to zero. Therefore, we can conclude that convergence of the proposed primal dual IPMs by using the new parameterization of the central trajectory is faster than classical primal dual IPMs.

6.2 Implementation Details

It is obvious that a good optimization software requires more than just a good optimization algorithm. Efficiency and stability of the linear algebra, initialization, termination, parameter selection are important issues to consider for a successful piece of optimization software. In this section, we describe in detail the steps we took to obtain a practical implementation of the proposed primal dual IPMs by using the new parameterization of the central trajectory. The most important issue is the linear algebra such as the problem of solving a large, sparse linear system at each iteration to find moving directions. Other issues that we discuss here include initialization of the algorithm, termination criteria, parameter selections. The modified and original primal dual IPM algorithms are implemented for both LO and QO problems by using MATLAB.

Termination Criteria: We used the duality gap as stopping criteria since convergence has been achieved when the duality gap becomes close to zero with given accuracy ε. Specifically,

$$\|c^T x - b^T y\| = \varepsilon. \tag{6.8}$$

Initialization of the Algorithm: One of the basic assumptions for our implementation is that we are given a feasible interior solution. Some initial solution finding techniques are explained in [44].

Parameter Selection: In the proposed primal dual IPMs by using the new parameterization of the central trajectory, we have some parameters such as $\varepsilon, \rho, \sigma, h$ that we need to select before execution. ε is the accuracy parameter for the termination criteria. In our implementation, we choose $\varepsilon = 0.0001$ and ρ is a parameter $0 < \rho < 1$ that we used it to define new points as in 5.27 and 5.56. In our implementation, we choose $\rho = 0.65$ where h is a parameter that defines step length in 5.19 and 5.43. In our implementation, we choose $h = 0.1$. σ is a parameter $0 < \sigma < 1$ that defines the reduction rate of μ at the original primal dual IPMs as in 3.19. We choose $\sigma = 0.5$.

Linear Algebra Issues: All of the matrices and vectors of our implementation used are predefined MATLAB functions. For advanced implementation of the algorithms advanced data structures can be considered. For each step of the proposed primal dual IPM, the most important task is to solve the system of equations for the new moving directions (dx, dy, dz). This consumes most of the computational tasks. There are two major computational tasks. First, we have to form the $AZ^{-1}XA^T$, where $Z^{-1}X$is a diagonal matrix and changes at each iteration. Second, we must solve the following system of equations

$$(AZ^{-1}XA^T)dy = AZ^{-1}Xd_D + d_P - AZ^{-1}d, \tag{6.9}$$

where $AZ^{-1}XA^T$ is symmetric and positive definite. Solving the system of equations could be done in various ways such as QR factorization, Cholesky factorization or preconditioned conjugate gradient. In our implementation MATLAB functions are used to deal with this problem.

Next, we discuss the ill-conditioning problem. We first define the condition of a matrix and then describe the ill-conditioning associated with proposed algorithm.

The condition of a problem can be defined as a measure that reflects the sensitivity of the exact solution to changes in the problem. Let $f(x)$ denote the exact solution of the given problem for variable x. If small changes in x lead to small changes in $f(x)$, the problem is *well-conditioned.* However, if small change lead to large ones, the problem is *ill-conditioned.* In a similar way, the condition of a matrix can be defined as follows [29].

Let A be a non singular matrix and consider the linear system $Ax = b$. Then, the exact solution is $x = A^{-1}b$. Notice that, the core work at each iteration of our proposed primal dual IPM algorithm is similar here. We find the direction vectors by solving a linear system at each iteration. The condition of a non singular matrix A is defined as

$$cond(A) = \|A^{-1}\|\|A\|.$$

Note that, $I = A^{-1}A$ and by using Cauchy-Schwarz inequality we get $1 = \|A^{-1}A\| \leq \|A^{-1}\|\|A\|$. We conclude with the inequality, $cond(A) \geq 1$.

A well-conditioned matrix has a condition number of order unity. An ill-conditioned matrix has a condition number much larger than unity. The condition of a matrix can be interpreted in terms of the closeness of the matrix to singularity. Informally speaking, an ill-conditioned matrix is near singular.

Here, we are interested in for the case where A is symmetric. More specifically, $AZ^{-1}XA^T$ is a symmetric matrix and we update and compute this matrix at each iteration. If a given matrix is symmetric, then the condition of a matrix can be defined as the ratio between the largest and the smallest eigenvalues as follows

$$cond(A) = \frac{|\lambda_{max}|}{|\lambda_{min}|},$$

where λ_{max} is the largest eigenvalue of A and λ_{min} is the smallest eigenvalue of A. The matrix is ill-conditioned if this ratio is very large. This usually happens when the smallest eigenvalues become near zero or when the largest eigenvalue goes to infinity.

Ill-conditioning is often observed during the final stages of a primal dual algorithm, when the elements of the diagonal matrix $AZ^{-1}XA^T$ take on both very large and very small values.

6.3 Computational Complexity Details

Computational complexity, originating from the interactions between computer science and optimization, is one of the major research areas that have revolutionized the approach for solving optimization problems and for analyzing their inherent difficulty. Computation within an iteration is obviously an important factor in the complexity of an algorithm. The types of computations within an iteration change little from one IPM algorithm to another, so that it is usually the number of iterations that distinguishes algorithms.

The number of iterations depends on the stopping criteria. Since, primal dual IPMs use the duality gap as a stopping criterion, the number of iterations is a function of the duality gap reduction. In general, the duality gap reduction at each iteration depends on the dimension n in a certain way that can be described as follows:

$$gap(k) = {z^k}^T x^k = n\mu^k \leq \varepsilon, \tag{6.10}$$

where

$$\mu^k = \mu^0 F(n)^k \tag{6.11}$$

and F is a function that defines the update rule for the barrier parameter. In order to find a bound for the number of iterations k for primal dual IPMs, the following inequality has to be satisfied

$$n\mu^0 F(n)^k \leq \varepsilon. \tag{6.12}$$

Kojima, Mizuno and Yoshise [38] showed that their primal dual IPM algorithm takes no more than $O(nL)$ iterations where L is the input size for the given problem.

The best known bound on the number of iterations for an IPM algorithm is $O(\sqrt{n}L)$ as presented by Renegar [56] and Monteiro and Adler [48, 49]. Monteiro and Adler's algorithm is a primal dual IPM where the duality gap reduction is a function of $\sqrt{n}$. Other techniques for bounding the number of iterations in path following algorithms are studied by Vaidya and Atkinson [67].

In this study, we proposed a modified primal dual IPM algorithm by using a new parameterization of the central trajectory. Each iterations of the proposed algorithm involves the inversion of a $n \times n$ matrix which can be done in at most $O(n^3)$ arithmetic operations. Furthermore, one can approximate the matrix to be inverted so that each iteration can be executed in the average of $O(n^{2.5})$ arithmetic operations. Since we have faster convergence results than the original primal dual IPMs, we can claim that the number of iterations of the proposed algorithm will be lower than the original primal dual IPMs. In other words, our proposed primal dual IPM algorithm takes less than $O(nL)$ iterations. Finally, our algorithm overall requires less than $O(n^{3.5}L)$ arithmetic iterations.

6.4 Stability Details

In the proposed primal dual IPM algorithm by using a new parameterization of the central trajectory, the rate of change of the barrier parameter μ is defined as a continuous dynamical system. In this section, we study the stability of the dynamical system and DAE system that resulted while parameterizing the central trajectory. Late in the nineteen century, Lyapunov a Russian mathematician, developed an approach to the stability analysis of dynamical systems. The main and unique feature of the approach is that only the form of differential or difference equations needs to be known, not their solutions. The method, called *Lyapunov's method*, requires evaluation of a function the so-called Lyapunov function. Evaluation of this function alone allows for the stability of the system to be proven.

Let us consider the autonomous system described with a system of n first-order differential equations:

$$\frac{dx}{dt} = f(x), x \in \Re^n. \tag{6.13}$$

We assume that the equations have been written so that $x^* = 0$ is an equilibrium point that satisfies the system of equations $f(x^*) = 0$. We formulate a condition for the equilibrium $x^* = 0$ to be asymptotically stable. This means that the state vector goes to zero as time goes to infinity. A function, $E(x)$, is said to be a *Lyapunov function*, if there exists an equilibrium state, x^*, such that the following three conditions are satisfied.

1. $E(x)$ is continuous with respect to all components of x,

2. $E(x)$ is positive definite. That is, $E(x^*) = 0$, and $E(x) > 0$ for $x \neq x^*$.

3. $\frac{dE(x)}{dt}$ is negative semi-definite. That is, the function is decreasing with time.

The function $E(x)$ is not unique, rather many different Lyapunov functions can be found for a given system. If at least one function is known that meets all the above conditions, the given system is asymptotically stable.

The dynamical system that resulted while parameterizing the central trajectory defined in 5.10 and 5.18 is as follows

$$\begin{aligned}\frac{dx}{dt} &= \frac{d\mu}{dt}\left[I - Z^{-1}XA^T(AZ^{-1}XA^T)^{-1}A\right]Z^{-1}e \\ \frac{dy}{dt} &= -\frac{d\mu}{dt}(AZ^{-1}XA^T)^{-1}AZ^{-1}e \\ \frac{dz}{dt} &= \frac{d\mu}{dt}A^T(AZ^{-1}XA^T)^{-1}AZ^{-1}e,\end{aligned} \tag{6.14}$$

where

$$\frac{d\mu}{dt} = \sum_{i=1}^{n} \ln x_i. \tag{6.15}$$

The stability of our dynamical system can be established by defining a function $E(x)$ and proving that $E(x)$ is a Lyapunov function as follows. Let us consider the problem P_μ in 5.4

$$\theta(\mu) = \inf\left\{c^T x - \mu\sum_{i=1}^{n}\ln x_i \quad \text{such that} \quad Ax = b, x > 0\right\}. \tag{6.16}$$

Let $E(\mu) = \theta(\mu) - \theta(\mu^*)$ where μ^* minimizes $\theta(\mu)$, for all $\mu \geq 0$. Now, we need to show that all conditions given above are satisfied by $E(\mu)$. Note that, $\theta(\mu)$ in 6.16 is strictly convex since both the objective function is strictly convex and the constraints are convex. Therefore, $E(\mu)$ is positive definite. From the previous section for the convergence results, we know that x, y, z, and μ are continuous in μ and t. Next we need to show that $E(\mu)$ is decreasing in time. Again, we know that the solution points x, y, z, and μ to the DAE system in 5.11 with initial points solve the KKT conditions for the given problem. Therefore, it follows from equations 5.6 and 5.9 that

$$\frac{dE(\mu)}{dt} = \frac{d\theta(\mu)}{d\mu}\frac{d\mu}{dt} = -\left[\frac{d\theta(\mu)}{d\mu}\right]^2 < 0. \tag{6.17}$$

This proves that $E(\mu)$ is a Lyapunov function, which implies that $E(\mu)$ decreases monotonically until a stable state is reached in which case neither the Lyapunov function nor the state changes further. Therefore stable points are the minimizers of $\theta(\mu)$. In other words, stable points are the optimal solutions of the given problem [79].

Next, we study the stability of the following DAE system described in 5.11.

$$\begin{aligned} \frac{d\mu}{dt} &= \sum_{i=1}^{n} \ln x_i \\ Ax - b &= 0 \\ A^T y + z &= c \\ XZe - \mu e &= 0. \end{aligned} \tag{6.18}$$

Note that this is an index-1 DAE as proved in Chapter 5. Ascher and Petzold [3] presented that if the all of the following conditions are satisfied by the index-1 DAE system, then the DAE system is stable. Specifically, for a linear index-1 problem, if

1. it can be transformed (without differentiations) into a semi-explicit system, and from there to an ODE by eliminating the algebraic variables,
2. the transformations are all suitably well conditioned,
3. the obtained ODE problem is stable,

then the index-1 DAE problem is also stable in the usual sense.

Next we need to show that the DAE system in 6.18 satisfies all of the above conditions. It is obvious that the DAE system is in a semi-explicit form as in equation

4.3. Note that during this formulation, the DAE system resulted in this semi-explicit form naturally. Semi-explicit DAE transformed to an ODE are shown in Chapter 5. So far, the first condition of the above conditions is satisfied. The resulting ODE after this transformation are shown in 5.16 as follows

$$\frac{dv}{dt} = J^{-1}(v) f(v, \mu, t). \tag{6.19}$$

Note that the Jacobian that defines the resulting ODE is the same as the Jacobian of the primal dual IPMs. By the assumption of nonsingularity of the Jacobian of primal dual IPMs, our transformation is well conditioned. That proves that the second condition is also satisfied. The resulting ODE is stable as we showed previously in this section by defining a Lyapunov function. That concludes that all above conditions are satisfied. Therefore we can state that the DAE system described in 6.18 is stable.

Chapter 7

Computational Results

In this chapter, we report some computational results for the proposed parameterization of the central trajectory for LO and convex QO problems and applications to support vector machines (SVMs) for the classification problem. The modified and classical primal dual IPM algorithms are implemented for both LO and QO problems by using MATLAB and codes are given in Appendix A. Numerical experiments are conducted by using some test problems to demonstrate the behavior and performance analysis of the new approach to the parameterization of the central trajectory. Iteration numbers for the proposed parameterization are compared with the classical parameterization. Note that our purpose here is to show graphically the functional and operational characteristics of the proposed trajectory and to illustrate the computational aspects of the methodological work, we have chosen some rather simple small size text-book examples. Corresponding parameter selections and other implementation issues are presented in Chapter 6.

7.1 LO Test Problems

To demonstrate the behavior and performance analysis of both of the trajectories the following LO test problems are used. Each problem is solved four times by using different initial solutions, same parameters and same stopping criteria by using primal dual IPM by the classical parameterization and modified primal dual IPM by the proposed parameterization. By doing so, we can make a fair comparison for the behavior and performance analysis of both of the trajectories. After solving each problem, some experiments are performed to analyze the central trajectory, the duality gap reductions and the barrier parameter μ reductions. Since our purpose here is to show graphically the functional and operational characteristics of the proposed trajectory, we choose some rather simple small size examples.

Problem LO1: We choose the following LO problem presented in [7].

$$
\begin{aligned}
\textit{LO1:} \quad & \min \quad -x_1 - 3x_2 \\
& \text{subject to} \\
& -x_1 + 2x_2 \leq 6 \\
& x_1 + x_2 \leq 5 \\
& x_1, x_2 \geq 0,
\end{aligned}
\tag{7.1}
$$

This problem is solved by using the following initial solutions.

$x^{01} = [1\ 1\ 5\ 3]^T$, $y^{01} = [0\ 0]^T$, and $z^{01} = 1/x^{01}$

$x^{02} = [1\ 2\ 3\ 2]^T$, $y^{02} = [0\ 0]^T$, and $z^{02} = 1/x^{02}$

$x^{03} = [2\ 1\ 6\ 2]^T$, $y^{03} = [0\ 0]^T$, and $z^{03} = 1/x^{03}$

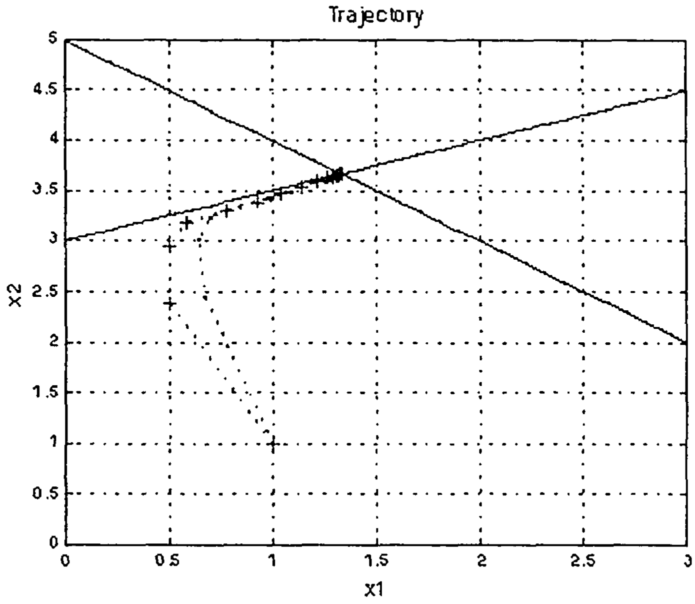

Figure 7.1: Central trajectory of problem LO1

$x^{04} = [3\ 1.5\ 7.5\ 0.5]^T$, $y^{04} = [0\ 0]^T$, and $z^{04} = 1/x^{04}$

The optimal solution is $x^* = [1.3333\ 3.6667]^T$, and the corresponding objective value is -9.6667. The number of iterations to converge to the optimal solution for these initial solutions are given in the Table 7.1. Clearly, the proposed parameterization has better solutions over the classical parameterization for this problem. Central trajectories of both modified and classical primal dual IPMs for the first initial point are shown in Figure 7.1. From that figure, we can see the different trajectories.

The duality gap reductions of both approaches for the first initial point are shown are shown in Figure 7.2. We see that the duality gap reductions of the proposed trajectory are faster than the classical trajectory. This result is consistent with the convergence results in Chapter 6. The μ^k values of both approaches converges to zero

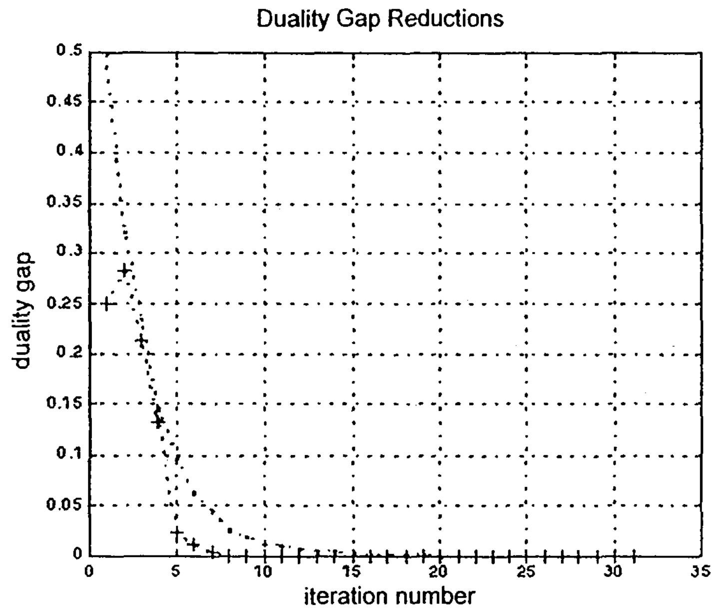

Figure 7.2: Duality gap reductions of problem LO1

as time increases. This is shown for the first initial point in Figure 7.3. We can see that, starting from the initial point μ_0 decreases to zero monotonically and rapidly for the proposed trajectory.

Problem LO2: We choose the following LO problem used in [72].

$$
\begin{aligned}
\textit{LO2:} \quad & \min \quad -2x_1 - 3.5x_2 \\
& \text{subject to} \\
& -x_1 + 4x_2 \leq 1 \\
& 2x_1 + 3x_2 \leq 3.5 \\
& 2x_1 + x_2 \leq 3 \\
& x_1, x_2 \geq 0,
\end{aligned} \tag{7.2}
$$

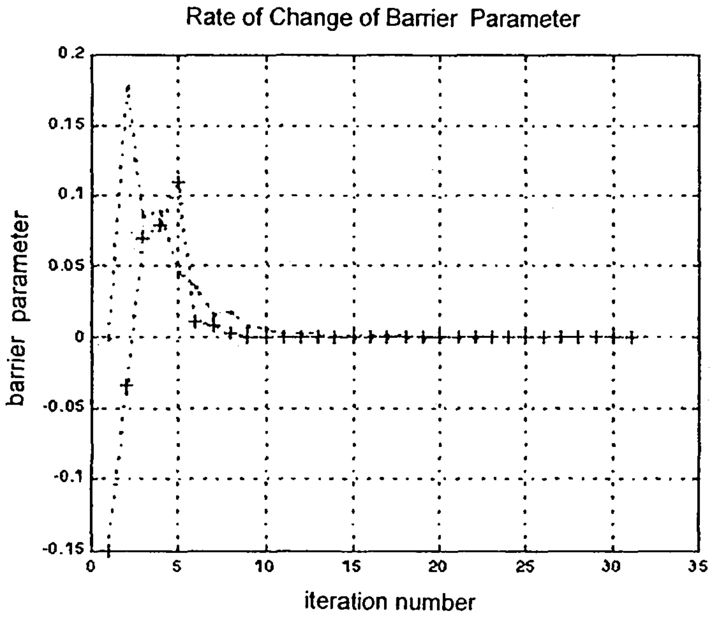

Figure 7.3: μ^k of problem LO1

This problem is solved by using the following initial solutions.

$x^{01} = [0.8\ 0.2\ 1\ 1.3\ 1.2]^T$, $y^{01} = [0\ 0\ 0]^T$, and $z^{01} = 1/x^{01}$

$x^{02} = [0.5\ 0.3\ 0.3\ 1.6\ 1.7]^T$, $y^{02} = [0\ 0\ 0]^T$, and $z^{02} = 1/x^{02}$

$x^{03} = [0.6\ 0.2\ 0.8\ 1.7\ 1.6]^T$, $y^{03} = [0\ 0\ 0]^T$, and $z^{03} = 1/x^{03}$

$x^{04} = [0.4\ 0.2\ 0.6\ 2.1\ 2]^T$, $y^{04} = [0\ 0\ 0]^T$, and $z^{04} = 1/x^{04}$

The optimal solution of this problem is $x^* = [1\ 0.5]^T$, and the corresponding objective value is -3.75. The number of iterations to converge to the optimal solution for these initial solutions are given in the Table 7.1. Clearly, modified primal dual IPM performs equally same over primal dual IPM for this problem for the given initial solutions. Central trajectories of both modified and classical primal dual IPMs for the first initial point are shown in Figure 7.4. From that figure, one can see that

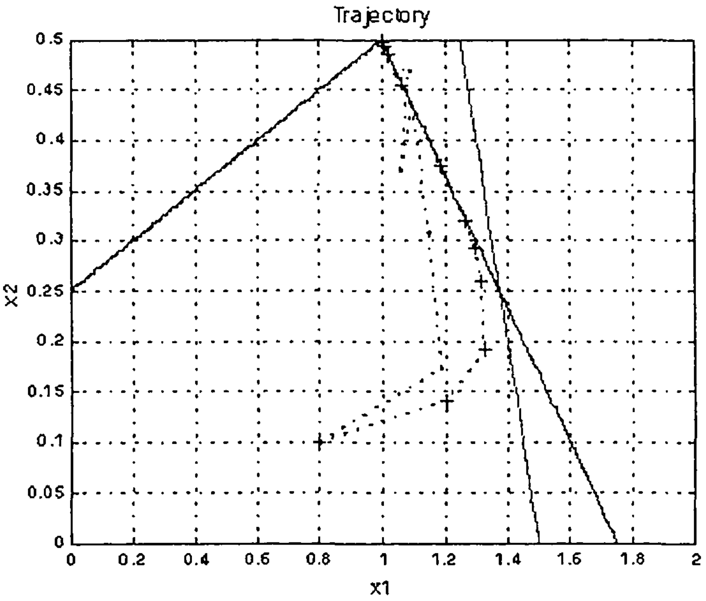

Figure 7.4: Central trajectory of problem LO2

both methods follow different trajectories. Although, classical trajectory reaches to the optimal solution almost same iteration numbers with the proposed trajectory, it can be seen from the duality gap reductions that we have consistent results as in the previous problem LO1.

The duality gap reductions of both approaches for the first initial point are shown in Figure 7.5. We see that the duality gap reductions of the proposed trajectory are faster than the classical one. The μ^k values of both approaches for the first initial point are shown in Figure 7.6.

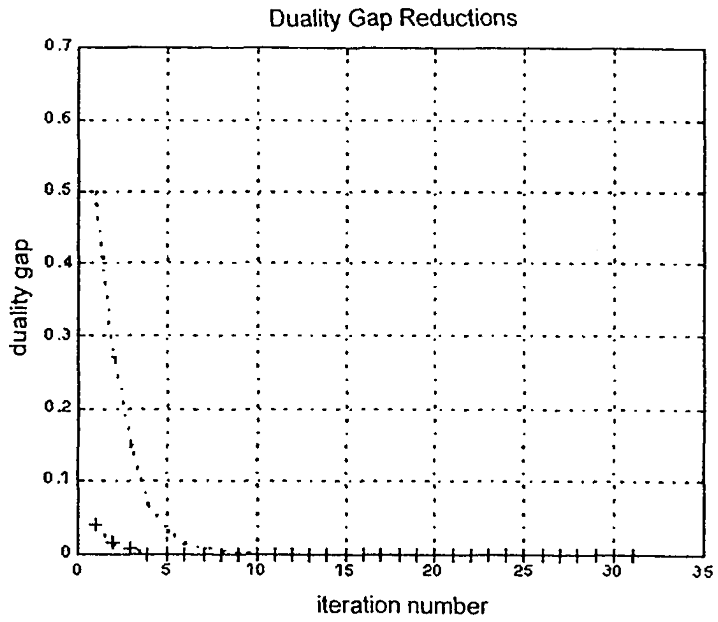

Figure 7.5: Duality gap reductions of problem LO2

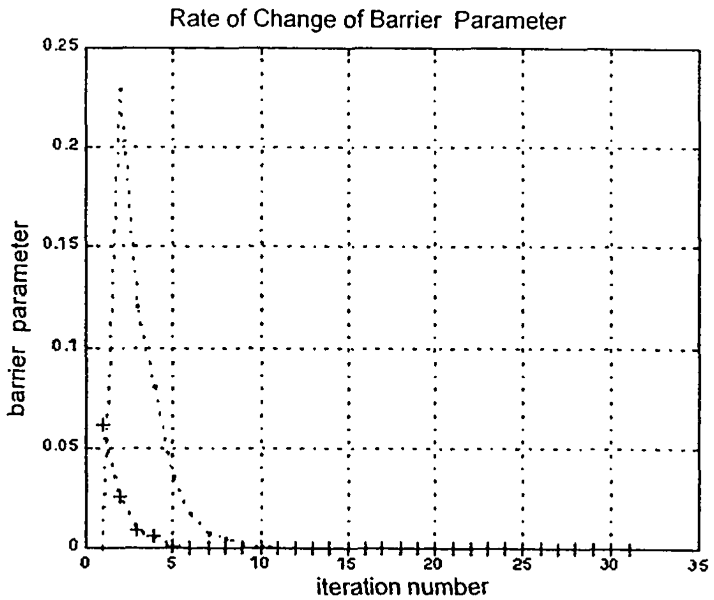

Figure 7.6: μ^k of Problem LO2

Problem LO3: We choose another simple LO problem from [7].

$$\begin{aligned} LO3: \quad & \min \quad 4x_1 - x_3 + x_4 \\ & \text{subject to} \\ & x_1 + x_2 + x_3 + x_4 = 1 \\ & -2x_1 + 2x_2 + x_3 - x_4 = 0 \\ & x_1, x_2, x_3, x_4 \geq 0, \end{aligned} \tag{7.3}$$

This problem is solved by using the following initial solutions.

$x^{01} = [0.1\ 0.1\ 0.4\ 0.4]^T$, $y^{01} = [0\ 0]^T$, and $z^{01} = 1/x^{01}$

$x^{02} = [0.25\ 0.25\ 0.25\ 0.25]^T$, $y^{02} = [0\ 0]^T$, and $z^{02} = 1/x^{02}$

$x^{03} = [0.1\ 0.3\ 0.1\ 0.5]^T$, $y^{03} = [0\ 0]^T$, and $z^{03} = 1/x^{03}$

$x^{04} = [0.2\ 0.2\ 0.3\ 0.3]^T$, $y^{04} = [0\ 0]^T$, and $z^{04} = 1/x^{04}$

The optimal solution of this problem is $x^* = [0\ 0\ 0.5\ 0.5]^T$, and the corresponding objective value is 0. The number of iterations to converge to the optimal solution for these initial solutions are given in the Table 7.1. Clearly, modified primal dual IPM performs equally same over primal dual IPM for this problem for the given initial solutions. The duality gap reductions of both approaches for the first initial point are shown in Figure 7.7. We see that the duality gap reductions of the proposed trajectory are faster than the classical one. The μ^k values of both approaches for the first initial point are shown in Figure 7.8.

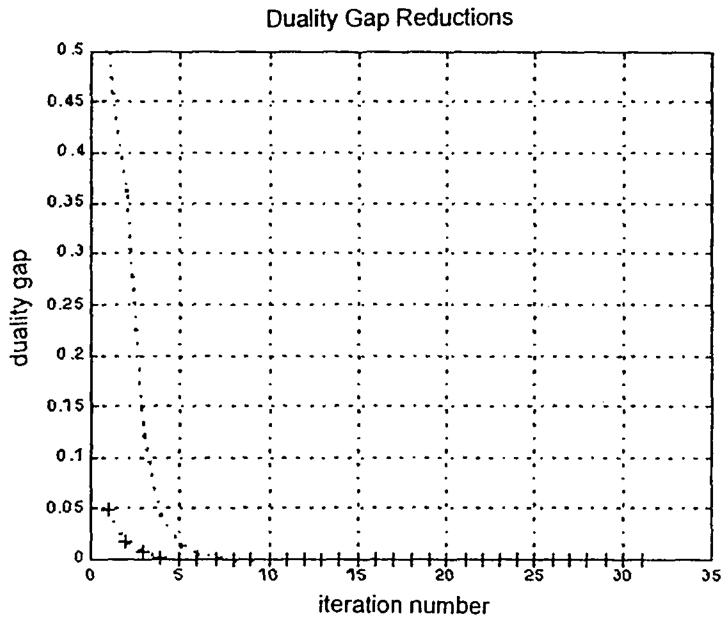

Figure 7.7: Duality gap reductions of problem LO3

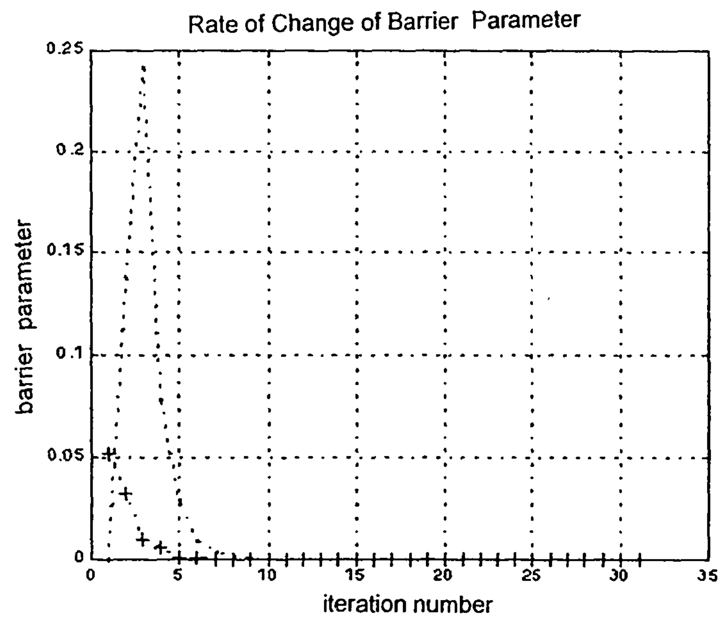

Figure 7.8: μ^k of problem LO3

Problem LO4: We choose another LO test problem presented in [73].

$$
\begin{aligned}
\textit{LO4:} \quad & \min \quad 2.4x_1 + 1.6x_2 + 4.2x_3 + 5.2x_4 + 2.4x_5 \\
& \text{subject to} \\
& \qquad -4.3x_1 + 5.3x_2 + 1.6x_3 + 0.5x_4 - 2.1x_5 = 12.5 \\
& \qquad 7.2x_1 - 2.6x_2 + 2.4x_3 + 1.6x_4 + 2.9x_5 = 7.2 \\
& \qquad 1.3x_1 - 1.2x_2 + 2.5x_3 + 4.1x_4 - 2.7x_5 = 6.3 \\
& \qquad x_1, x_2, x_3, x_4, x_5 \geq 0,
\end{aligned} \tag{7.4}
$$

This problem is solved by using the following initial solutions.

$x^{01} = [0.1295\ 2.2951\ 3.2307\ 1\ 1]^T$, $y^{01} = [0\ 0\ 0]^T$, and $z^{01} = 1/x^{01}$

$x^{02} = [0.6170\ 1.9585\ 3.0832\ 0.1\ 0.1]^T$, $y^{02} = [0\ 0\ 0]^T$, and $z^{02} = 1/x^{02}$

$x^{03} = [0.7309\ 2.7750\ 1.2720\ 2\ 1]^T$, $y^{03} = [0\ 0\ 0]^T$, and $z^{03} = 1/x^{03}$

$x^{04} = [0.7432\ 2.2114\ 2.4591\ 0.5\ 0.2]^T$, $y^{04} = [0\ 0\ 0]^T$, and $z^{04} = 1/x^{04}$

The optimal solution of this problem is $x^* = [0.6711\ 1.9631\ 3.1133]^T$, and the corresponding objective value is 17.8277. The number of iterations to converge to the optimal solution for these initial solutions are given in the Table 7.1. Clearly, modified primal dual IPM performs equally same over primal dual IPM for this problem for the given initial solutions.

Problem LO5 - Global Routing Problem: This problem is a simple global routing problem using a LO formulation presented in [70]. In this problem, we want to connect 3 modules by a wire using horizontal and vertical segments. A measure which is used to obtain global routings is the minimization of wirelength. Using wirelength measure, we have the following LO representation of the global routing problem.

$$
\begin{aligned}
\textit{LO5:} \qquad & \min \quad -4x_1 - 3x_2 - 3x_3 - 4x_4 - x_5 - 3x_6 \\
& \text{subject to} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (7.5) \\
& x_1 + x_2 + x_3 + x_4 \leq 1 \\
& x_5 + x_6 \leq 1 \\
& x_1 + x_4 + x_6 \leq c_1 \\
& x_1 + x_2 + x_3 + x_5 \leq c_2 \\
& x_1 + x_4 + x_6 \leq c_3 \\
& x_3 + x_4 + x_6 \leq c_4 \\
& x_2 \leq c_5 \\
& x_1 + x_3 + x_4 \leq c_6 \\
& x_2 \leq c_7 \\
& x_1, x_2, x_3, x_4, x_5, x_6 \geq 0.
\end{aligned}
$$

This problem is solved by using the following initial solutions.

$x^{01} = [1\ 3\ 2\ 2\ 1\ 3\ 2\ 2\ 1\ 3\ 2\ 2\ 1\ 3\ 2]^T$, $y^{01} = [0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]^T$, and $z^{01} = 1/x^{01}$

$x^{02} = [1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1]^T$, $y^{02} = [0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]^T$, and $z^{02} = 1/x^{02}$

$x^{03} = [1\ 2\ 3\ 4\ 1\ 1\ 3\ 1\ 1\ 3\ 1\ 4\ 3\ 2\ 1]^T$, $y^{03} = [0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]^T$, and $z^{03} = 1/x^{03}$

$x^{04} = [1\ 2\ 1\ 4\ 1\ 1\ 2\ 1\ 4\ 1\ 1\ 2\ 1\ 4\ 1]^T$, $y^{04} = [0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]^T$, and $z^{04} = 1/x^{04}$

The values c_i are the number of available tracks in each edge. If $c_i = 1$ for all i, then the optimal solution of the problem is $x^* = [0\ 1\ 0\ 0\ 0\ 1]^T$, and the corresponding objective value is -6. The number of iterations to converge to the optimal solution for these initial solutions are given in the Table 7.1. Clearly, modified primal dual IPM performs equally same over primal dual IPM for this problem for the given initial solutions.

Table 7.1: Summary of the computations for LO test problems

	n	m	PD1	MD1	PD2	MPD2	PD3	MPD3	PD4	MPD4
LO1	4	2	33	20	14	13	32	20	13	16
LO2	5	3	12	13	12	12	12	13	13	13
LO3	4	2	11	11	13	16	13	14	12	11
LO4	5	3	13	16	14	14	13	15	12	13
LO5	15	9	16	14	15	13	16	18	17	13

7.2 Convex QO Test Problems

To demonstrate the behavior and performance analysis of both of the trajectories the following convex QO problems are used. Since our purpose here was to show graphically the functional and operational characteristics of the proposed trajectory, we have chosen some rather simple small size text-book examples. Each problem is solved four times by using different initial solutions, same parameters and same stopping criteria by using primal dual IPM by the classical parameterization and modified primal dual IPM by the proposed parameterization. By doing so, we can make a fair comparison for the behavior and performance analysis of the trajectories.

Problem QO1: The following convex QO problem is presented in [55].

$$\begin{array}{lll} QO1: & \min & -6x_1 + 2x_1^2 - 2x_1x_2 + 2x_2^2 \\ & \text{subject to} & \\ & & x_1 + x_2 \leq 2 \\ & & x_1, x_2 \geq 0, \end{array} \tag{7.6}$$

This problem is solved by using the following initial solutions.

$x^{01} = [1\ 0.5\ 0.5]^T$, $y^{01} = [0]^T$, and $z^{01} = 1/x^{01}$

$x^{02} = [0.3\ 1\ 0.7]^T$, $y^{02} = [0]^T$, and $z^{02} = 1/x^{02}$

$x^{03} = [0.486\ 0.681\ 0.833]^T$, $y^{03} = [0]^T$, and $z^{03} = 1/x^{03}$

$x^{04} = [0.8335\ 0.2484\ 0.9181]^T$, $y^{04} = [0]^T$, and $z^{04} = 1/x^{04}$

The optimal solution of this problem is $x^* = [1.5\ 0.5]^T$ and the corresponding objective value is -5.5. The number of iterations to converge to the optimal solution for these initial solutions are given in the Table 7.2. Clearly, modified primal dual IPM performs equally same over primal dual IPM for this problem for the given initial solutions.

Problem QO2: The following convex QO problem is presented in [60].

$$\begin{array}{lll} QO2: & \min & -4x_1 - 6x_2 + 2x_1^2 + 2x_1x_2 + 2x_2^2 \\ & \text{subject to} & \\ & & x_1 + 2x_2 \leq 2 \\ & & x_1, x_2 \geq 0, \end{array} \tag{7.7}$$

This problem is solved by using the following initial solutions.

$x^{01} = [0.2\ 0.5\ 0.8]^T$, $y^{01} = [0]^T$, and $z^{01} = 1/x^{01}$

$x^{02} = [1\ 1\ 1]^T$, $y^{02} = [0]^T$, and $z^{02} = 1/x^{02}$

$x^{03} = [5\ 10\ 1]^T$, $y^{03} = [0]^T$, and $z^{03} = 1/x^{03}$

$x^{04} = [1\ 2\ 3]^T$, $y^{04} = [0]^T$, and $z^{04} = 1/x^{04}$

The optimal solution of this problem is $x^* = [0.3333\ 0.8333]^T$ and the corresponding objective value is -4.1667. The number of iterations to converge to the optimal solution for these initial solutions are given in the Table 7.2. Clearly, modified primal dual IPM performs equally same over primal dual IPM for this problem for the given initial solutions.

Problem QO3: We choose another simple QO problem from [7].

$$\begin{aligned} QO3: \quad & \min \quad -2x_1 - 6x_2 + x_1^2 - 2x_1x_2 + 2x_2^2 \\ & \text{subject to} \\ & \qquad x_1 + x_2 \leq 2 \\ & \qquad -x_1 + 2x_2 \leq 2 \\ & \qquad x_1, x_2 \geq 0. \end{aligned} \tag{7.8}$$

This problem is solved by using the following initial solutions.

$x^{01} = [5\ 0.1\ 0.5\ 2]^T$, $y^{01} = [0\ 0]^T$, and $z^{01} = 1/x^{01}$

$x^{02} = [1\ 1\ 1\ 1]^T$, $y^{02} = [0\ 0]^T$, and $z^{02} = 1/x^{02}$

$x^{03} = [5\ 10\ 1\ 2]^T$, $y^{03} = [0\ 0]^T$, and $z^{03} = 1/x^{03}$

$x^{04} = [1\ 2\ 3\ 4]^T$, $y^{04} = [0\ 0]^T$, and $z^{04} = 1/x^{04}$

The optimal solution of this problem is $x^* = [0.8\ 1.2]^T$ and the corresponding objective value is -7.2. The number of iterations to converge to the optimal solution for these initial solutions are given in the Table 7.2. Clearly, modified primal dual IPM performs equally same over primal dual IPM for this problem for the given initial solutions.

Problem QO4: We choose another QO problem from [73].

$$
\begin{aligned}
QO4: \quad & \min \quad 1.5x_1^2 - x_1x_2 + x_1x_3 - 2x_1x_4 + 2x_2^2 + 2x_2x_3 \\
& \qquad +2.5x_3^2 + x_3x_4 + 3x_4^2 - 6x_1 + 15x_2 + 9x_3 + 4x_4 \\
& \text{subject to} \\
& \qquad x_1 + 2x_2 + 4x_3 + 5x_4 = 12 \\
& \qquad 3x_1 - 2x_2 - x_3 + 2x_4 = 8 \\
& \qquad 2x_1 - 3x_2 + x_3 - 4x_4 = 6 \\
& \qquad x_1, x_2, x_3, x_4 \geq 0,
\end{aligned}
\tag{7.9}
$$

This problem is solved by using the following initial solutions.

$x^{01} = [1\ 5\ 5\ 2]^T$, $y^{01} = [0\ 0\ 0]^T$, and $z^{01} = 1/x^{01}$

$x^{02} = [10\ 1\ 10\ 1]^T$, $y^{02} = [0\ 0\ 0]^T$, and $z^{02} = 1/x^{02}$

$x^{03} = [5\ 10\ 1\ 2]^T$, $y^{03} = [0\ 0\ 0]^T$, and $z^{03} = 1/x^{03}$

$x^{04} = [1\ 2\ 8\ 1]^T$, $y^{04} = [0\ 0\ 0]^T$, and $z^{04} = 1/x^{04}$

The optimal solution of this problem is $x^* = [2.967\ 0\ 1.7363\ 0.4176]^T$ and the corresponding objective value is 24.1578. The number of iterations to converge to the optimal solution for these initial solutions are given in the Table 7.2. Clearly, modified primal dual IPM under performs over primal dual IPM for this problem for the given initial solutions.

Table 7.2: Summary of the computations for QO test problems

	n	m	PD1	MD1	PD2	MPD2	PD3	MPD3	PD4	MPD4
QO1	3	1	11	11	12	13	11	12	12	12
QO2	3	1	12	13	10	11	11	13	12	18
QO3	4	2	14	15	12	13	14	16	13	16
QO4	4	3	14	18	16	18	15	20	14	18

7.3 Applications to Support Vector Machines (SVMs) for Classification

Support vector machines (SVMs) [71] have recently attracted much attention in optimization and learning theory. As a new tool for solving problems in machine learning, they are based on quadratic optimization approaches. Since the problem is convex, there are no local minima and various optimization algorithms will be able to identify the optimal solution. The name support vector is derived from those points in the input space which touch ("support") the decision function. An overview of SVMs can be found in [14]. Learning machines can be implemented in two different ways as classification and regression respectively. In classification problems, the output parameter y_i is a categorical variable that indicates to which class a given input vector x_i belongs to. For classification of linearly nonseparable data, the decision function is given by

$$D(x) = \sum_{i=1}^{n} \alpha_i^* y_i H(x_i, x), \tag{7.10}$$

where H is a kernel function [14]. The parameters α_i^* define the solution for the following quadratic optimization problem.

$$\max \quad Q(\alpha) = \sum_{i=1}^{n} \alpha_i - 0.5 \sum_{j,i=1}^{n} \alpha_i \alpha_j y_i y_j H(x_i x_j)$$

$$\text{subject to} \qquad \qquad \qquad (7.11)$$

$$\begin{aligned} \sum_{i=1}^{n} y_i \alpha_i &= 0 \\ \alpha_i &\leq C/n \\ \alpha_i &\geq 0. \end{aligned}$$

The training data $(x_i x_j)$, inner product kernel H, and regularization parameter C are given.

In this section, the proposed parameterization of the central trajectory is used to solve the resulting SVM quadratic optimization problem for the XOR problem. The XOR problem for SVMs can be defined as follows; find an optimal separating hyperplane that classifies the following data set given in Table 7.3 without error. It is not possible to solve this problem with a linear decision boundary since the data

Table 7.3: XOR data

index i	x_1	x_2	y
1	1	1	1
2	1	-1	-1
3	-1	1	-1
4	-1	-1	1

are nonlinear nonseparable. However, a polynomial decision boundary of order 2 can separate these data [16]. The inner product kernel for polynomials of order 2 is

$$H(x, x') = [(x.x') + 1]^2. \tag{7.12}$$

This expression corresponds to the set of basis functions

$$z = (1, \sqrt{2}x_1, \sqrt{2}x_2, \sqrt{2}x_1x_2, x_1^2, x_2^2), \tag{7.13}$$

where x_1, x_2 correspond to the two input space coordinates. This vector z is a point in a six-dimensional feature space. To determine the decision boundary in this space, we must solve a convex optimization problem. Depending on the norm we are using it can be either LO or convex QO problem.

Linear Optimization Problem: We define the XOR problem for SVMs by using the LO problems as follows.

$$\begin{aligned} \min \quad & \sum_{i=1}^{4} \alpha_i + C \sum_{i=1}^{4} Z_i \\ & \text{subject to} \\ y_i \left(\sum_{j=1}^{4} y_j \alpha_j [H(x_j, x_i)] + b \right) + z_i \quad & \geq \quad 1 \\ \alpha_i, z_i \quad & \geq \quad 0. \end{aligned} \tag{7.14}$$

The inner product kernel is represented as a 4×4 matrix H with elements H_{ij} computed by using 7.12 and the given data. Specifically,

$$H = \begin{bmatrix} 9 & 1 & 1 & 1 \\ 1 & 9 & 1 & 1 \\ 1 & 1 & 9 & 1 \\ 1 & 1 & 1 & 9 \end{bmatrix} \tag{7.15}$$

By substituting the values of kernel function $H(x_j, x_i)$ and y_i output values from Table 7.3, then we can rewrite this problem as follows:

$$
\begin{aligned}
\min \quad & \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + c(z_1 + z_2 + z_3 + z_4) \\
\text{subject to} \quad & \\
& \begin{aligned}
-[9\alpha_1 - \alpha_2 - \alpha_3 + \alpha_4 + b] - z_1 + 1 &\leq 0 \\
[\alpha_1 - 9\alpha_2 - \alpha_3 + \alpha_4 + b] - z_2 + 1 &\leq 0 \\
[\alpha_1 - \alpha_2 - 9\alpha_3 + \alpha_4 + b] - z_3 + 1 &\leq 0 \\
-[\alpha_1 - \alpha_2 - \alpha_3 + 9\alpha_4 + b] - z_4 + 1 &\leq 0 \\
\alpha_i, z_i &\geq 0.
\end{aligned}
\end{aligned}
\tag{7.16}
$$

This problem is solved by using primal dual IPM and modified primal dual IPM by using the following initial solutions.

$x^{01} = [0.08\ 0.1\ 0.14\ 0.15\ 1\ 0.7\ 0.8\ 0.1\ 1\ 1.5\ 0.6\ 0.2]^T$, $y^{01} = [0\ 0\ 0\ 0]^T$, and $z^{01} = 1/x^{01}$

$x^{02} = [1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1]^T$, $y^{02} = [0\ 0\ 0\ 0]^T$, and $z^{02} = 1/x^{02}$

$x^{03} = [1\ 2\ 3\ 4\ 5\ 6\ 6\ 5\ 4\ 3\ 2\ 1]^T$, $y^{03} = [0\ 0\ 0\ 0]^T$, and $z^{03} = 1/x^{03}$

$x^{04} = [1\ 2\ 1\ 0.4\ 1\ 3\ 4\ 1\ 0.1\ 1\ 3\ 0.2]^T$ $y^{04} = [0\ 0\ 0\ 0]^T$ and $z^{04} = 1/x^{04}$

The optimal solution to this linear optimization problem for C=10000 is $\alpha_i = 0.125$ and the corresponding objective value is 0.5. The number of iterations to converge to the optimal solution for these initial solutions are given in the Table 7.4. Clearly, modified primal dual IPM performs better over primal dual IPM for this problem for the given initial solutions.

Quadratic Optimization Problem: Now, we define the SVM QO for the XOR problem for SVM by using QO problem as follows.

$$\begin{aligned} \min \quad & -\alpha_1 - \alpha_2 - \alpha_3 - \alpha_4 + 0.5 \sum_{i,j=1}^{4} \alpha_i \alpha_j y_i y_j H_{ij} \\ \text{subject to} \quad & \\ & \alpha_1 - \alpha_2 - \alpha_3 + \alpha_4 = 0 \\ & \alpha_i \geq 0. \end{aligned} \tag{7.17}$$

The inner product kernel is represented as a 4×4 matrix H with elements H_{ij} computed as LO case and the corresponding data is same as 7.15. Therefore, objective function of the quadratic problem can be written as follows.

$$\min \quad -\alpha_1 - \alpha_2 - \alpha_3 - \alpha_4 + 0.5 \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \end{bmatrix} \begin{bmatrix} 9 & -1 & -1 & 1 \\ -1 & 9 & 1 & -1 \\ -1 & 1 & 9 & -1 \\ 1 & -1 & -1 & 9 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix}$$

This problem is solved by using primal dual IPM and modified primal dual IPM by using the following four initial solutions.

$x^{01} = [1\ 0.2\ 5\ 1]^T$, $y^{01} = [0]^T$, and $z^{01} = 1/x^{01}$

$x^{02} = [1\ 1\ 1\ 1]^T$, $y^{02} = [0]^T$, and $z^{02} = 1/x^{02}$

$x^{03} = [1\ 2\ 3\ 4]^T$, $y^{03} = [0]^T$, and $z^{03} = 1/x^{03}$

$x^{04} = [2\ 3\ 0.2\ 4]^T$, $y^{04} = [0]^T$, and $z^{04} = 1/x^{04}$

The optimal solution to this convex quadratic optimization problem for is $\alpha_i = 0.125$ indicating that all four data points are support vectors and the corresponding objective value is -0.25. Computational results for these initial solutions are given in the

Table 7.4. The number of iterations to converge to the optimal solution for these initial solutions are given in the Table 7.4. Clearly, modified primal dual IPM performs better over primal dual IPM for this problem for the given initial solutions.

Table 7.4: Summary of the computations for SVMs for pattern recognition

	n	m	PD1	MD1	PD2	MPD2	PD3	MPD3	PD4	MPD4
LO	12	4	47	20	30	20	58	27	74	21
QO	4	1	17	11	12	12	21	12	20	12

Chapter 8

Summary, Conclusions and Recommendations

8.1 Summary

In this work, a new approach to the parameterization of the central trajectory for primal dual IPMs is proposed. A continuous dynamical system that describes the rate of change of the barrier parameter of the central trajectory is considered. Instead of parameterizing the central trajectory by the barrier parameter, it is parameterized by the time. Therefore the central trajectory is described through continuous dynamical system. Specifically, a new update rule based on the solution of an ODE for the barrier parameter of the primal dual IPMs is presented. The resulting ODE combined with the first order Karush-Kuhn-Tucker conditions, which are algebraic equations, are called differential algebraic equations (DAEs). By solving DAEs, we follow approximately the central trajectory of the primal dual IPMs. By doing so, we find an optimal solution to the given problem.

The proposed parameterization of the central trajectory is investigated both for LO and convex QO problems. Primal dual IPMs for LO and convex QO problems are modified by using the proposed approach to the parameterization of central trajectory. We proposed two new primal dual IPM algorithms based on these modifications.

The proposed parameterization of the central trajectory is studied in detail for convergence, implementation, computational complexity and stability issues. We proved the convergence of the proposed algorithms and showed that they converge faster than original primal dual IPMs. Stability of DAEs are also proven that the resulting DAEs are stable. Numerical experiments are conducted by using some test problems to demonstrate the behavior and performance analysis of the new approach to the parameterization of the central trajectory. Computational results are consistent with the theoretical results.

8.2 Recommendations for Future Research

In this study, we considered only linear and convex quadratic optimization problems for the proposed primal dual IPMs by using the new parameterization of the central trajectory. The proposed parameterization of the central trajectory for primal dual IPMs can be extended to the general nonlinear optimization problems and positive semidefinite optimization.

Since our purpose here was to show graphically the functional and operational characteristics of the proposed trajectory, we have chosen some rather simple small size text-book examples. Therefore an experimentation with bigger size and real world test problems can be performed to get better feeling about the performance of the proposed trajectory in large scale setting.

In this study, to solve ODE we use the backward Euler's method. Different ODE solvers like one-step, multi-step, Runge-Kutta, predictor-corrector techniques can be be tested to solve the resulting DAE. An investigation for the best ODE solver to solve the DAE for mathematical optimization problems may result better performance for the algorithms.

Bibliography

[1] I. Adler and R.D.C. Monteiro. Convergence and boundary behavior of the projective scaling trajectories for linear programming. In J.C. Lagarias and M.J. Todd, editors, *Mathematical Developments Arising from Linear Programming, Contemporary Mathematics, Vol. 114*, pages 189–211. American Mathematical Society, 1990.

[2] K.M. Anstreicher. A monotonic projective algorithm for fractional linear programming. *Algorithmica*, 1:483–498, 1986.

[3] U. M. Ascher and L. R. Petzold. *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations.* Society for Industrial and Applied Mathematics, SIAM, Philadelphia, Pennsylvania, USA, 1998.

[4] E.R. Barnes. A variation of Karmarkar's algorithm for linear programming problems. *Mathematical Programming*, 36:174–182, 1986.

[5] 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*, 314:499–526, 1989.

[6] 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*, 314:527–581, 1989.

[7] M.S. Bazaraa, H.D. Sherali, and C.M. Shetty. *Nonlinear Programming: Theory and Algorithms.* John Wiley and Sons, New York, 1993.

[8] D.P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods.* Academic Press, New York, 1982.

[9] M.G. Breitfeld and D.F. Shanno. Computational experience with penalty-barrier methods for nonlinear programming. RUTCOR Research Report, RRR 17-93, August 1993, Rutgers Center for Operations Research, Rutgers University, New Brunswick, New Jersey, 1993.

[10] M.G. Breitfeld and D.F. Shanno. A globally convergent penalty-barrier algorithm for nonlinear programming and its computational performance. RUTCOR Research Report, RRR 12-94, April 1994, Rutgers Center for Operations Research, Rutgers University, New Brunswick, New Jersey, 1994.

[11] M.G. Breitfeld and D.F. Shanno. Preliminary computational experience with modified log-barrier functions for large-scale nonlinear programming. In W.W. Hager, D.W. Hearn, and P.M. Pardalos, editors, *Large-Scale Optimization: State of the Art*, pages 45–67. Kluwer Academic Publishers, 1994.

[12] K.E. Brenan, S.L. Campbell, and L.R. Petzold. *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations.* North-Holland, Amsterdam, Holland, 1989.

[13] A.A. Brown and M.C. Bartholomew-Biggs. Some effective methods for unconstrained optimization based on the solution of systems of ordinary differential equations. *Journal of Optimization Theory and Applications*, 62(2):211–224, 1989.

[14] C.J.C. Burges. A tutorial on support vector machines for pattern classification. *Data Mining and Knowledge Discovery*, 2(2):121–167, 1998.

[15] A. M. Cervantesa, A. Wachtera, R. H. Tutuncu, and L. T. Biegler. A reduced space interior point strategy for optimization of differential algebraic systems. *Computers and Chemical Engineering*, 24(1):39–51, 2000.

[16] V. Cherkassky and F. Mulier. *Learning from Data: Concepts, Theory, and Methods.* John Wiley and Sons, New York, USA, 1998.

[17] N. Couellan. *Primal-Dual Techniques for Nonlinear Programming and Applications to Artificial Neural Network Training.* PhD thesis, School of Industrial Engineering, University of Oklahoma, Norman, Oklahoma, 1997.

[18] D. DenHertog. *Interior Point Approach to Linear, Quadratic and Convex Programming: Algorithm and Complexity.* Kluwer Academic Publishers, Dordrecht, Netherlands, 1994.

[19] I.I. Dikin. Iterative solution of problems of linear and quadratic programming. *Soviet Mathematics Doklady*, 8:674–675, 1967.

[20] A.S. El-Bakry, R.A. Tapia, T. Tsuchiya, and Y. Zhang. On the formulation and theory of the Newton interior-point method for nonlinear programming. *Journal of Optimization Theory and Applications*, 89(3):507–541, 1996.

[21] S.C. Fang and S. Puthenpura. *Linear Optimization and Extensions: Theory and Algorithms.* Prentice-Hall, Englewood Cliffs, New Jersey, 1993.

[22] L.E. Faybusovich. Dynamical systems which solve optimization problems with linear constraints. *IMA Journal of Mathematical Control and Information*, 8:135–149, 1991.

[23] L.E. Faybusovich. Hamiltonian structure of dynamical systems which solve linear programming problems. *Physica D*, 53:217–232, 1991.

[24] A.V. Fiacco and G.P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques.* John Wiley and Sons, New York, 1968.

[25] K.R. Frisch. The logarithmic potential method for convex programming. Institute of Economics, University of Oslo, Oslo, Norway, 1955.

[26] D.M. Gay. A variant of Karmarkar's linear programming algorithm for problems in standard form. *Mathematical Programming*, 37:81–90, 1987.

[27] C.W. Gear. Simultaneous numerical solution of differential-algebraic equations. *IEEE Transactions on Circuit Theory*, TC-18(1):89–95, 1971.

[28] C.W. Gear and L.R. Petzold. ODE methods for the solution of differential/algebraic systems. *SIAM Journal of Numerical Analysis*, 21(4):716–728, 1984.

[29] P. E. Gill, W. Murray, , and M. H. Wright. *Numerical Linear Algebra and Optimization.* Addison-Wesley, Reading, Massachusetts, USA, 1991.

[30] V. Gopal and L. T. Biegler. A successive linear programming approach for initialization and reinitialization after discontinuities of differential-algebraic equations. *SIAM Journal on Scientific Computing*, 20(2):447–467, 1999.

[31] E. Hairer, C. Lubich, and M. Roche. *The numerical solution of differential-algebraic systems by* Runge-Kutta *methods*, volume 1409 of *Lecture Notes in Mathematics.* Springer-Verlag, Heidelberg, Berlin, Germany, 1989.

[32] E. Hairer and G. Wanner. *Solving ordinary differential equations* II: *Stiff and differential-algebraic problems*, volume 14 of *Computational Mathematics.* Springer-Verlag, Heidelberg, Berlin, Germany, 1996.

[33] A.C. Hindmarsh and L.R. Petzold. Algorithms and software for ordinary differential equations and differential-algebraic equations part II: Higher order methods and software packages. *Computers in Physics*, 9(2):148–155, 1995.

[34] A.C. Hindmarsh and L.R. Petzold. Algorithms and software for ordinary differential equations and differential-algebraic equations part I: Euler methods and error estimation. *Computers in Physics*, 9(1):34-41, 1995.

[35] P. Huard. Resolution of mathematical programming with nonlinear constraints by the method of centers. In J. Abadie, editor, *Nonlinear Programming*, pages 207–219. North-Holland, 1967.

[36] W. Kaplan. *Advanced Calculus.* Addison-Wesley, New York, USA, 1991.

[37] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.

[38] M. Kojima, S. Mizuno, and A. Yoshise. A primal-dual interior point algorithm for linear programming. In N. Megiddo, editor, *Progress in Mathematical Programming: Interior-Point and Related Methods*, pages 29–48. Springer-Verlag, 1989.

[39] P. Lotstedt and L. Petzold. Numerical solution of nonlinear differential equations with algebraic constraints I: convergence results for backward differentiation formulas. *Mathematics of Computation*, 46(174):491–516, 1986.

[40] I.J. Lustig, R.E. Marsten, and D.F. Shanno. On implementing Mehrotra's predictor-corrector interior point method for linear programming. *SIAM Journal on Optimization*, 2:435–449, 1992.

[41] I.J. Lustig, R.E. Marsten, and D.F. Shanno. Interior point method for linear programming. *ORSA Journal on Computing*, 6(1):1–14, 1994.

[42] T. Maly and L.R. Petzold. Numerical methods and software for sensitivity analysis of differential-algebraic systems. 1995.

[43] R.E. Marsten, R. Subramanian, M.J. Saltzman, I.J. Lustig, and D.F. Shanno. Interior point methods for linear programming: Just call Newton, Lagrange, and Fiacco and McCormick! *Interfaces*, 20(4):105–116, 1990.

[44] K.A. McShane, C.L. Monma, and D. Shanno. An implementation of a primal-dual interior point method for linear programming. *ORSA Journal on Computing*, 1(2):70–83, 1989.

[45] N. Megiddo. Pathways to the optimal solution in linear programming. In N. Megiddo, editor, *Progress in Mathematical Programming: Interior-Point and Related Methods*, pages 131–158. Springer-Verlag, 1989.

[46] N. Megiddo and M. Shub. Boundry behavior of interior point algorithms in linear programming. *Mathematics of Operational Research*, 14:97–146, 1989.

[47] S. Mehrotra and J. Sun. An interior point method for solving smooth convex programs based on Newton's method. In J.C. Lagarias and M.J. Todd, editors, *Mathematical Developments Arising from Linear Programming, Contemporary Mathematics, Vol. 114*, pages 265–284. American Mathematical Society, 1990.

[48] R.D.C. Monteiro and I. Adler. Interior path following primal-dual algorithms. Part I: linear programming. *Mathematical Programming*, 44:27–41, 1989.

[49] R.D.C. Monteiro and I. Adler. Interior path following primal-dual algorithms. Part II: convex quadratic programming. *Mathematical Programming*, 44:43–66, 1989.

[50] S.G. Nash, R. Polyak, and A. Sofer. A numerical comparison of barrier and modified barrier functions for large scale bound constrained optimization. In W.W. Hager, D.W. Hearn, and P.M. Pardalos, editors, *Large-Scale Optimization: State of the Art*, pages 319–338. Kluwer Academic Publishers, 1994.

[51] Y. Nesterov and A. Nemirovskii. *Interior Point Polynomial Methods in Convex Programming*, volume 13 of *SIAM Studies in Applied Mathematics*. SIAM, Philadelphia, PA, USA, 1994.

[52] M. A. Nunez and R. M. Freund. Condition measures and properties of the central trajectory of a linear program. *Mathematical Programming*, 83:1–28, 1998.

[53] L. Petzold. Differential/algebraic equations are not ODEs. *SIAM Journal of Scientific Statistical Computing*, 3(3):367–384, 1982.

[54] R. Polyak. Modified barrier functions (theory and methods). *Mathematical Programming*, 54:177–222, 1992.

[55] A. Ravindran, D.T. Phillips, and J.J. Solberg. *Operations Research: Principles and Practice*. John Wiley and Sons, New York, 1987.

[56] J. Renegar. A polynomial-time algorithm based on Newton's method for linear programming. *Mathematical Programming*, 40:59–93, 1988.

[57] J.G. Renfro. *Computational Studies in the Optimization of Systems Described by Differential/Algebraic Equations*. PhD thesis, Department of Chemical Engineering, University of Houston, University Park, Houston, Texas, 1986.

[58] W.C. Rheinboldt. Differential/algebraic systems as differential equations on manifolds. *Mathematics of Computation*, 43(168):473–482, 1984.

[59] C. Roos and J.P. Vial. A polynomial method of approximate centers for linear programming. *Mathematical Programming*, 54:295–305, 1992.

[60] A.H. Taha. *Operations Research: An Introduction.* MacMillan Publishing Company, New York, 1992.

[61] M.J. Todd and B.P. Burrell. An extension of Karmarkar's algorithm for linear programming using dual variables. *Algorithmica*, 1:409–424, 1986.

[62] T. B. Trafalis and S. Kasap. Differential algebraic equations in primal dual interior point methods for linear programming problems. In C.H. Dagli, M. Akay, A.L. Buczak, O. Ersoy, B.R. Fernandez, and J. Ghosh, editors, *Intelligent Engineering Systems Through Artificial Neural Networks*, volume 8, pages 323–328. ASME Press, 1998.

[63] T.B. Trafalis and T.A. Tutunji. A stochastic logarithmic barrier function method for neural network training with bounded weights. In C.H. Dagli, M. Akay, C.L.P. Chen, B. Fernandez, and J. Ghosh, editors, *Intelligent Engineering Systems Through Artificial Neural Networks, Vol. 5*, pages 167–172. ASME Press, 1995.

[64] T.B. Trafalis and T.A. Tutunji. Deterministic and stochastic logarithmic barrier function methods for neural network training. In A. Migdalas et al., editor, *Par-*

allel Computing in Optimization, pages 529–574. Kluwer Academic Publishers, 1997.

[65] L. Tuncel and M.J. Todd. On the interplay among entropy, variable metrics and potential functions in interior-point algorithms. 1995.

[66] T.A. Tutunji. *Logarithmic Barrier Functions and Newton-Type Methods with Applications to Neural Network Training*. PhD thesis, School of Industrial Engineering, University of Oklahoma, Norman, Oklahoma, 1996.

[67] P. M. Vaidya and D. S. Atkinson. A technique for bounding the number of iterations in path following algorithms. In Panos M. Pardalos, editor, *Complexity in Numerical Optimization*, pages 462–489. World Scientific, 1993.

[68] R.J. Vanderbei, M.S. Meketon, and B.A. Freedman. A modification of Karmarkar's linear programming algorithm. *Algorithmica*, 1:395–407, 1986.

[69] R.J. Vanderbei and D.F. Shanno. An interior point algorithm for nonconvex nonlinear programming. SOR-97-21, Statistics and Operations Research, Princeton University, New Jersey, 1997.

[70] A. Vannelli. Teaching large-scale optimization by an interior point approach. *IEEE Transactions on Education*, 36(1):204–209, 1993.

[71] V. Vapnik. *The Nature of Statistical Learning Theory*. Springer Verlag, New York, 1995.

[72] J. Wang. Deterministic neural networks for combinatorial optimization. In O.M. Omidvar, editor, *Progress in Neural Networks*, pages 319–340. Ablex, 1994.

[73] J. Wang. Recurrent neural networks for optimization. In C.H. Chen, editor, *Fuzzy Logic and Neural Network Handbook*, pages 4.1–4.35. McGraw-Hill, 1996.

[74] C. Witzgall, P.T. Boggs, and P.O. Domich. On the convergence behavior of trajectories for linear programming. In J.C. Lagarias and M.J. Todd, editors, *Mathematical Developments Arising from Linear Programming, Contemporary Mathematics, Vol. 114*, pages 161–187. American Mathematical Society, 1990.

[75] H. Wolkowicz, R. Saigal, and L. Vanderberghe. *Handbook of Semidefinite Programming - Theory, Algorithms, and Applications.* Kluwer Academic Publishers, Boston, MA, USA, 2000.

[76] J.S. Wright. *Primal-Dual Interior-Point Methods.* Society for Industrial and Applied Mathematics, SIAM, Philadelphia, Pennsylvania, USA, 1997.

[77] M. Xiong, J. Wang, and P. Wang. Differential-algebraic approach to linear programming. *Journal of Optimization Theory and Applications*, 114(2):443–470, 2002.

[78] Y. Ye. An extension of Karmarkar's algorithm and the trust region method for quadratic programming. In N. Megiddo, editor, *Progress in Mathematical Programming: Interior-Point and Related Methods*, pages 49–64. Springer-Verlag, 1989.

[79] W.I. Zangwill and C.B. Garcia. *Pathways to Solutions, Fixed Points, and Equilibria.* Prentice-Hall, Inc., Englewoods Cliffs, New Jersey, USA, 1981.

Appendix A

MATLAB Codes

Main Program file named as "mpdlp.m" for modified primal dual IPMs for linear optimization problems. It calls some functions that are given after the main program file.

```
clear session;

clear all;

lo1;

%lo2;

%lo3;

%lo4;

%lo5;

%xor;

i=0; m0=0.1;

while abs(c'*x0-b'*y0)> 0.0001
```

```
i=i+1;

trajectory(i,1)=x0(1);

trajectory(i,2)=x0(2);

X=diag(x0);

Z=diag(z0);

t0=b-A*x0;

u0=c-A'*y0-z0;

v0=m0*e-X*Z*e+h*dm(x0,n);

dy=inv(A*inv(Z)*X*A')*(A*inv(Z)*X*u0+t0-A*inv(Z)*v0);

dz=u0-A'*dy;

dx=inv(Z)*(v0-X*dz);

s=1;

t=1;

stepx(s)=s;

stepz(t)=t;

for j=1:n

  if dx(j) < 0

     stepx(s)=-x0(j)/dx(j);

     s=s+1;

  end

  if dz(j) < 0

     stepz(t)=-z0(j)/dz(j);
```

```
        t=t+1;
      end
    end
    sp=alpha*min(stepx);
    sd=alpha*min(stepz);
    x0=x0+sp*dx;
    y0=y0+sd*dy;
    z0=z0+sd*dz;
    m1=m0+h*dm(x0,n);
    while m1 < 0
       h=h/2;
       m1=m0+h*dm(x0,n);
    end
    m0=m1;
    tram0(i)=m0;
end
x0
i
c'*x0
```

The following function calculates the rate of change of the barrier parameter μ when the main program needs it.

```
function y=dm(x,n)
```

```
sum1=0;
for i=1:n
   sum1=sum1+log(x(i));
end
y=sum1;
```

An example of an input file named as "lo4.m" that contains problem data, initial solution and parameter values.

```
n=5; m=3;
A=[-4.3 5.3 1.6 0.5 -2.1;7.2 -2.6 2.4 1.6 2.9;1.3 -1.2 2.5 4.1 -2.7];
b=[12.5;7.2;6.3];
c=[2.4;1.6;4.2;5.2;2.4];
%x0=[0.1295;2.2951;3.2307;1;1];
%x0=[0.617;1.9585;3.0832;0.1;0.1];
%x0=[0.7309;2.775;1.272;2;1];
x0=[0.7432;2.2114;2.4591;0.5;0.2];
y0=[0;0;0];
z0=[1/x0(1);1/x0(2);1/x0(3);1/x0(4);1/x0(5)];
e=[1;1;1;1;1];
alpha=0.65; sigma=0.5; h=0.1;
```

Main Program file named as "pdlp.m" for classical primal dual IPMs for linear optimization problems.

```
clear session;
```

```
clear all;
%lo1;
lo2;
%lo3;
%lo4;
%lo5;
%xor;
i=0;
m0=0.5;
while abs(c'*x0-b'*y0)> 0.0001
   i=i+1;
   trajectory(i,1)=x0(1);
   trajectory(i,2)=x0(2);
   X=diag(x0);
   Z=diag(z0);
   m0=sigma*(x0'*z0/n);
   pathm0(i)=m0;
   t0=b-A*x0;
   u0=c-A'*y0-z0;
   v0=m0*e-X*Z*e;
   dy=inv(A*inv(Z)*X*A')*(A*inv(Z)*X*u0+t0-A*inv(Z)*v0);
   dz=u0-A'*dy;
```

```
  dx=inv(Z)*(v0-X*dz);
  s=1;
  t=1;
  stepx(s)=s;
  stepz(t)=t;
  for j=1:n
    if dx(j) < 0
       stepx(s)=-x0(j)/dx(j);
       s=s+1;
    end
    if dz(j) < 0
       stepz(t)=-z0(j)/dz(j);
       t=t+1;
    end
  end
  sp=alpha*min(stepx);
  sd=alpha*min(stepz);
  x0=x0+sp*dx;
  y0=y0+sd*dy;
  z0=z0+sd*dz;
end
x0
```

```
i

c'*x0
```

Main Program file named as "mqpo.m" for modified primal dual IPMs for convex quadratic optimization problems.

```
clear session;
clear all;
%qp1;
%qp2;
%qp3;
qp4;
%qxor;
i=0;
m0=0.1;
while abs(c'*x0-b'*y0+x0'*Q*x0)> 0.0001
   i=i+1;
   X=diag(x0);
   Z=diag(z0);
   dP=b-A*x0;
   dD=Q*x0+c-A'*y0-z0;
   dW=m0*e-X*Z*e+h*dm(x0,n);
   ZZ=Z+X*Q;
   dy=inv(A*inv(ZZ)*X*A')*(A*inv(ZZ)*X*dD+dP-A*inv(ZZ)*dW);
```

```
dx=inv(ZZ)*(dW-X*dD+X*A'*dy);

KK=dW-Z*dx;

for kk=1:n

   dzz(kk)=KK(kk)/x0(kk);

end

dz=dzz';

s=1;

t=1;

stepx(s)=s;

stepz(t)=t;

for j=1:n

  if dx(j) < 0

     stepx(s)=-x0(j)/dx(j);

     s=s+1;

  end

  if dz(j) < 0

     stepz(t)=-z0(j)/dz(j);

     t=t+1;

  end

end

sp=alpha*min(stepx);

sd=alpha*min(stepz);
```

```
    x0=x0+sp*dx;
    y0=y0+sd*dy;
    z0=z0+sd*dz;
    m1=m0+h*dm(x0,n);
    while m1 < 0
       h=h/2;
       m1=m0+h*dm(x0,n);
    end
    m0=m1;
    tram0(i)=m0;
end
i
x0
c'*x0+0.5*x0'*Q*x0
x0'*z0
```

Main Program file named as "qpo.m" for classical primal dual IPMs for convex quadratic optimization problems.

```
clear session;
clear all;
%qp1;
qp2;
%qp3;
```

```
%qp4;
qxor;
i=0;
while abs(c'*x0-b'*y0+x0'*Q*x0)> 0.0001
   i=i+1;
   X=diag(x0);
   Z=diag(z0);
   m0=sigma*(x0'*z0/n);
   dP=b-A*x0;
   dD=Q*x0+c-A'*y0-z0;
   dW=m0*e-X*Z*e;
   ZZ=Z+X*Q;
   dy=inv(A*inv(ZZ)*X*A')*(A*inv(ZZ)*X*dD+dP-A*inv(ZZ)*dW);
   dx=inv(ZZ)*(dW-X*dD+X*A'*dy);
   KK=dW-Z*dx;
   for kk=1:n
      dzz(kk)=KK(kk)/x0(kk);
   end
   dz=dzz';
   s=1;
   t=1;
   stepx(s)=s;
```

```
   stepz(t)=t;
   for j=1:n
     if dx(j) < 0
        stepx(s)=-x0(j)/dx(j);
        s=s+1;
     end
     if dz(j) < 0
        stepz(t)=-z0(j)/dz(j);
        t=t+1;
     end
   end
   sp=alpha*min(stepx);
   sd=alpha*min(stepz);
   x0=x0+sp*dx;
   y0=y0+sd*dy;
   z0=z0+sd*dz;
end
i
x0
c'*x0+0.5*x0'*Q*x0
x0'*z0
```

An example of an input file named as "qp4.m" that contains problem data, initial solution and parameter values.

```
n=4; m=3;
A=[1 2 4 5;3 -2 -1 2;2 -3 1 -4];
b=[12;8;6];
Q=[3 -1 1 -2;-1 4 2 0; 1 2 5 1;-2 0 1 6];
c=[-6;15;9;4];
%x0=[1;5;5;2];
%x0=[10;1;10;1];
%x0=[5;10;1;2];
x0=[1;2;8;1];
y0=[0;0;0];
z0=[1/x0(1);1/x0(2);1/x0(3);1/x0(4)];
e=[1;1;1;1;1];
alpha=0.65; sigma=0.5; h=0.1;
```

The following file tracks both of the trajectories of the given problem.

```
clear session;
clear all;
lo3;
pdlp;
clear A X Z alpha b c dx dy dz e h i j m m0 n s sd;
clear sigma sp stepx stepz t t0 v0 u0 x0 y0 z0 deltam01 m1 pm01;
```

```
lo3;

mpdlp;

draw3a;
```

The following file tracks the duality gap reductions of the given problem.

```
clear session;

clear all;

lo3;

pdlp;

clear A X Z alpha b c dx dy dz e h i j m m0 n s sd trajectory1;

clear sigma sp stepx stepz t t0 v0 u0 x0 y0 z0 deltam01 m1;

lo3;

mpdlp;

drawb;
```

The following file tracks the rate of change of the barrier parameter of the given problem.

```
clear session;

clear all;

lo3;

pdlp;

clear A X Z alpha b c dx dy dz e h i j m m0 n s sd trajectory1;

clear sigma sp stepx stepz t t0 v0 u0 x0 y0 z0 pm01 m1;
```

```
lo3;

mpdlp;

drawc;
```

The following file named as "draw3a.m" draws a figure that shows the trajectories.

```
[x1,x2]=meshgrid(0:0.01:2,0:0.01:0.5);

y=-x1+4.*x2;

v=[1, 10];

contour(x1,x2,y,v)

grid;

hold

y=2.*x1+3.*x2;

v=[3.5 10];

contour(x1,x2,y,v)

y=2.*x1+x2;

v=[3 10];

contour(x1,x2,y,v)

plot(trajectory1(:,1),trajectory1(:,2),'k.:')

plot(trajectory2(:,1),trajectory2(:,2),'k+:')

xlabel('x1','FontSize',12)

ylabel('x2','FontSize',12)

title('Trajectory','FontSize',12)
```

The following file named as "drawb.m" draws a figure that shows the duality gap reduction.

```
plot(pm01(:),'r.:')

grid;

hold

plot(pm02(:),'k+:')

xlabel('iteration number','FontSize',12)

ylabel('duality gap','FontSize',12)

title('Duality Gap Reductions','FontSize',12)
```

The following file named as "drawc.m" draws a figure that shows the rate of change of barrier parameter.

```
plot(deltam01(:),'r.:')

grid;

hold

plot(deltam02(:),'k+:')

xlabel('iteration number','FontSize',12)

ylabel('barrier parameter','FontSize',12)

title('Rate of Change of Barrier Parameter','FontSize',12)
```