The Notion of the Quasicentral Path in Linear Programming*

Miguel Argáez Osvaldo Mendez Leticia Velázquez

Abstract

The notion of the central path plays an important role in the development of most primal-dual interior-point algorithms. In this work we prove that a related notion called the quasicentral path, introduced by Argáez and Tapia in nonlinear programming, while being a less restrictive notion it is sufficiently strong to guide the iterates towards a solution of the problem. We use a new merit function for advancing to the quasicentral path, and weighted neighborhoods as proximity measures of this central region. We prove global convergence theory, and present some numerical results that demonstrate the effectiveness of the algorithm.

Keywords: Interior-point methods, primal-dual methods, linear programming, Newton's method, and merit function.

1 Introduction

The area of linear programming has been extensively studied in the last decades obtaining several well-known theoretical and numerical results. A book by Wright [11] presents most of the theoretical advances in linear programming, and a paper by Bixby [4] gives a brief

---

*Drs. Argáez and Velázquez would like to acknowledge the support by the U.S. Department of the Army Grant No. DAAD19-01-1-0741, and the Department of Computational and Applied Mathematics at Rice University. The authors thank Dr. Richard A. Tapia for his reviews and helpful comments of this paper. Correspondence to: Leticia Veláquez, email:`leti@utep.edu`.

summary of the computational developments for solving real-world linear programs. In particular, the work of Karmarkar [7] is noted for its role in promoting primal-dual interior-point algorithms (See, for example Kojima, Mizuno, and Yoshise [8], Lustig, Marsten, and Shanno [9], El-Bakry, Tapia, Tsuchiya, and Zhang [6], and Zhang [12]). Such approaches are based on using a central region, called the central path, as a guide for obtaining approximate solutions. Here, we introduce a new methodology that is based on a different central region, and present a global convergence theory.

In this work, we carry over a globalization strategy, presented by Argáez and Tapia [2], from nonlinear programming to linear programming. This strategy consists of following a related notion of the central path, called quasicentral path, as a central region for guiding the iterates towards a solution of the problem. An important result is that the dual variable y is not needed, at least explicitly, to find a solution of the problem. Specifically we prove that if the initial point is chosen so that the norm of the dual conditions is less than or equal to the norm of the primal conditions, then the convergence behavior to zero of the dual conditions depends on the convergence behavior of the primal conditions. Therefore we can exclude the dual conditions of the central path, obtaining the notion of quasicentral path as a central region suitable for guiding the iterates to a solution of the problem.

This leads us to present a path-following algorithm that is set in the framework of the Kojima et. al [8] algorithm. The path-following algorithm that we are proposing begins with a linesearch Newton's method applied to the perturbed KKT conditions for a fixed value $\mu > 0$ until an iterate belongs to a specific weighted neighborhood of the quasicentral path. If a solution of the problem is not found, then the perturbation parameter μ is reduced, a new weighted neighborhood is defined, and the Newton linesearch procedure is repeated. In order to monitor progress to the quasicentral path we present a new merit function and as proximity measures to this region we use specific weighted neighborhoods. Some important global properties of the merit function are presented, including a brief comparative discussion between weighted and non-weighted neighborhoods.

Finally, a global convergence theory and numerical experimentation are presented. We emphasize that the numerical experimentation shows only that the proposed technique works as well as current techniques on small to medium size problems. Further research is needed to demonstrate its competitiveness for a class of large scale problems.

1.1 Outline

This paper is organized as follows. In Section 2 we present the formulation of the problem. In Section 3 we describe our path following algorithm. In Section 4 we present a new merit function and some global properties. In Section 5 we define weighted neighborhoods as measures of proximity to the quasicentral path. In Section 6 we present a global convergence theory for the algorithm. In Section 7 we present numerical experimentation. And finally in Section 8 we give some conclusions and suggested future work.

2 Problem Formulation

We consider the linear programming problem in the standard form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0, \end{array} \tag{1}$$

where c, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$ $(m < n)$, and A is full rank. This problem is called the primal problem. The dual problem associated with problem (1) can be written

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & A^T y + z = c \\ & z \geq 0, \end{array} \tag{2}$$

where $y \in \mathbb{R}^m$ and $z \in \mathbb{R}^n$.

A point (x, z) is said to be a positive point if $x > 0$ and $z > 0$. A point (x, y, z) is said to be an interior point for the primal and dual problems if (x, z) is a positive point.

The optimality conditions, known as the Karush-Kuhn-Tucker (KKT) conditions, for the primal and dual problems are

$$F(x,y,z) = \begin{pmatrix} Ax - b \\ A^T y + z - c \\ XZe \end{pmatrix} = 0, \qquad (x,z) \geq 0. \tag{3}$$

where $X = \operatorname{diag}(x),\ Z = \operatorname{diag}(z)$, and $e = (1, \ldots, 1)^T \in \mathbb{R}^n$.

For problems (1) and (2), we define the feasible set as

$$\mathcal{F} = \Big\{(x,y,z) \in \mathbb{R}^{n+m+n} :\ Ax = b,\ A^T y + z = c,\ (x,z) \geq 0\Big\},$$

and the strictly feasible set as

$$\mathcal{F}^o = \Big\{(x,y,z) \in \mathcal{F} :\ (x,z) > 0\Big\}.$$

The solution set is

$$\mathcal{S} = \Big\{(x^*,y^*,z^*) \in \mathcal{F} :\ X^* Z^* e = 0\Big\}.$$

This set is one of the faces of the polyhedral $\mathcal{F}$.

If $\mathcal{F}^o$ is not empty, then S is also not empty and bounded. All the points in the relative interior, $ri(S)$, are strictly complementarity solutions, i.e., $x_i^* + z_i^* > 0$ for $i = 1, ..., n$. And the zero-nonzero pattern of the points in $ri(S)$ is invariant. Therefore, for any $(x^*, y^*, z^*) \in ri(S)$ the following index sets $\mathcal{B} = \{i : x_i^* > 0, \text{ for } i = 1, 2, \ldots, n\}$ and $\mathcal{N} = \{i : z_i^* > 0, \text{ for } i = 1, 2, \ldots, n\}$ are independent of the choice of a solution in $ri(S)$. Moreover, by strict complementarity $\mathcal{B} \cup \mathcal{N} = \{1, 2, \ldots, n\}$ and $\mathcal{B} \cap \mathcal{N} = \emptyset$.

In particular, among the set of solutions in $ri(S)$ there is one solution, called the analytic center, and denoted by (x_c^*, y_c^*, z_c^*) , such that

$$(x_c^*, y_c^*, z_c^*) = \arg\max \prod_{i \in \mathcal{B}} x_i \prod_{j \in \mathcal{N}} z_j.$$

In some linear programming applications, the primary objective is to compute the analytic center; however, the primary objective of this work is to promote the notion of the quasicentral path for solving linear programming problems.

The analytic center is associated with the notion of central path. For $\mu > 0$, the central path is defined as the set of points (x, y, z) satisfying the following perturbed KKT conditions

$$F_\mu(x, y, z) = \begin{pmatrix} Ax - b \\ A^T y + z - c \\ XZe - \mu e \end{pmatrix} = 0, \qquad (x, z) \geq 0. \tag{4}$$

This system has a unique solution $(x(\mu), y(\mu), z(\mu))$ for each fixed μ. Therefore the set $\{(x(\mu), y(\mu), z(\mu)), \mu > 0\}$ defines a smooth curve called the central path. As μ converges to zero the central path runs through the strictly feasible set $\mathcal{F}^o$, keeping an adequate distance from the non-optimal faces of $\mathcal{F}$, and ending at the analytic center, i.e.,

$$(x(\mu), y(\mu), z(\mu)) \to (x_c^*, y_c^*, z_c^*) \ \text{ as } \ \mu \to 0.$$

This classical result is applied in linear programming for obtaining an optimal solution of the primal and dual problems simultaneously.

Even though the notion of the central path plays an important role in the primal-dual interior-point methodology, a related notion of this region called the quasicentral path can be considered also as a central region for calculating a solution of the primal-dual problem. Then, the principal objective in this paper is to promote the notion of the quasicentral path.

In the work by Argáez and Tapia [2], a related notion of the central path was considered in nonlinear programming. In the arena of linear programming this notion is defined as follows: The quasicentral path is defined as the set of points (x, z) satisfying the following relaxation of the perturbed KKT conditions

$$\hat{F}_\mu(x, z) = \begin{pmatrix} Ax - b \\ XZe - \mu e \end{pmatrix} = 0, \qquad (x, z) > 0, \tag{5}$$

parameterized by $\mu > 0$.

Remark 2.1. *It is worth noticing that the quasicentral path defines a variety instead of a path.*[1]

The next property shows that the quasicentral path is equivalent to the region of strictly feasible points for the primal problem. Therefore following this region as a central region is equivalent to being strictly feasible with respect to the primal problem.

Property 2.1. *For $\mu > 0$, the projection of the quasicentral path defined by (5) on the x-space, coincides with the set of strictly feasible points.*

Proof. Let x be a strictly feasible point. Define the vector

$$z = \mu x^{-1}.$$

Then it is easily checked that (x, z) is on the quasicentral path. Conversely, x is a strictly feasible point if (x, z) is on the quasicentral path. $\square$

The following discussion in Property 2.2 provides the motivation for the use of the quasicentral path as a central region.

Property 2.2. *If the initial point is not a feasible point and by applying a damped Newton's method, then the dual residual, e_d^k, converges to zero if and only if the primal residual, e_p^k, converges to zero.*

Proof. By applying a damped Newton's method to primal and dual residuals at an infeasible starting point, then

$$\begin{aligned} e_p^1 &= b - Ax_1 = (1 - \alpha_1)e_p^o, \quad \text{and} \\ e_d^1 &= c - A^T y_1 - z_1 = (1 - \alpha_1)e_d^o. \end{aligned}$$

Iteratively, we obtain

$$e_p^k = b - Ax_k = (1 - \alpha_k)e_p^{k-1} = \prod_{j=1}^{k}(1 - \alpha_j)e_p^o, \text{ and}$$

---

[1] Argaez and Tapia chose the name of quasicentral path due to the fact that one of the conditions of the central path is omitted. The authors are fully aware of the fact that they use the term "quasicentral path" to denote mathematically would be known as a variety. However, we choose to retain the already established terminology originally introduced by Argaez and Tapia [1, 2].

$$e_d^k = c - A^T y_k - z_k = (1-\alpha_k)e_d^{k-1} = \prod_{j=1}^{k}(1-\alpha_j)e_d^o.$$

The proof follows from the last two equations. □

This property shows that e_p^k and e_d^k converge to zero at the same rate, but not necessarily in the same number of Newton iterations. Nevertheless, if the initial point is chosen so that $\|e_d^o\| \le \|e_p^o\|$, then the above property shows that the convergence to zero of the dual conditions depends on the convergence to zero of the primal conditions. In other words, e_d^k is zero if e_p^k is zero. In this situation, then we can remove the dual conditions from the central path, and consider the quasicentral path as a central region to be followed for obtaining a solution of the primal and dual problems simultaneously.

3 A Path-Following Algorithm

We present a path-following algorithm that uses the quasicentral path as a central region to guide the iterates toward a solution of the problem. To make progress to this region we use a new merit function, and specific weighted neighborhoods (see Definition 5.1) as measures of proximity to the quasicentral path. We start with an initial positive point (x, z) such that the error of the dual conditions e_d^o be less than or equal to the error of the primal conditions e_p^o, i.e. $\|e_d^o\| \le \|e_p^o\|$.

The algorithm being proposed is a path-following version of the Kojima-Mizuno-Yoshise algorithm. We follow the same globalization philosophy as in Argáez-Tapia [2], which consists of excluding the dual variable y and the dual condition e_d for effects of global convergence.

Algorithm 1

Step 1. Consider an initial positive point (x, z) and $\mu > 0$. Set $e_d = c - z$, $e_p = b - Ax$, $e_c = \mu e - XZe$, such that $\|e_d\| \le \|e_p\|$.

Step 2. Newton step. Solve the linear system for $(\Delta x, \Delta y, \Delta z)$

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Z & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = \begin{pmatrix} e_p \\ e_d \\ e_c \end{pmatrix} \tag{6}$$

Step 3. Maintain x and z positive. Choose $\tau \in (0,1)$ and calculate $\tilde{\alpha} = \min(1, \tau\hat{\alpha})$ where

$$\hat{\alpha} = \frac{-1}{\min(X^{-1}\Delta x, Z^{-1}\Delta z)}.$$

Step 4. Sufficient decrease. Find $\alpha = (\frac{1}{2})^t \tilde{\alpha}$ where t is the smallest positive integer such that

$$\Phi_\mu(x + \alpha\Delta x, z + \alpha\Delta z) \leq \Phi_\mu(x, z) + 10^{-4}\alpha\nabla\Phi_\mu(x, z)^T(\Delta x, \Delta z).$$

Update $(x, z) = (x, z) + \alpha(\Delta x, \Delta z)$, $e_p = (1 - \alpha)e_p$, and $e_d = (1 - \alpha)e_d$.

Step 5. Proximity to the quasicentral path. Choose an $\gamma \in (0,1)$.

If $\Big(\|e_p\|^2 + \|(XZ)^{-1/2}(XZe - \mu e)\|^2 \Big) \leq \gamma\mu$, then go to Step 6.

else, set $e_c = \mu e - XZe$, and go to step 2.

Step 6. Stopping criteria.

If $\Big(2\|e_p\| + x^T z)/(1 + \|b\|\Big) < \epsilon$, then stop

else update μ, set $e_c = \mu e - XZe$, and go to Step 2.

Remark 3.1. *In Step 4, the updates e_p and e_d are explained by Property 2.2, and the merit function Φ_μ is presented in Definition 4.1.*

Remark 3.2. *Observe that for fixed $\mu > 0$, the Algorithm 1 applies a linesearch Newton's method to the perturbed KKT conditions until an iterate (x, z) satisfies the inequality given in Step 5. This part of the algorithm is called the inner loop. If the iterate is not a solution of the problem, then the parameter μ is reduced and the procedure is repeated. The sequence consisting of the iterates that satisfy the inequality in Step 5 is called the outer loop of the algorithm.*

4 Merit Function and Global Properties

In the Algorithm 1 only the variables x and z are taken into account. The variable y is considered only implicitly. We use the notation $\tilde{v} = (x, z)$ as opposed to the standard notation $v = (x, y, z)$, in which the three variables are displayed.

Now for $\mu > 0$, the Newton step at the positive point $\tilde{v} = (x, z)$ is defined by $\Delta\tilde{v} = (\Delta x, \Delta z)$ where Δx and Δz are obtained from (6).

The purpose in this section is to present a new merit function that it forces the Newton iterates to advance towards the quasicentral path.

Definition 4.1. *For $\mu > 0$, we define the function*

$$\begin{aligned} &\Phi_\mu : \mathbb{R}^{n+n}_{++} \to \mathbb{R} \\ &\Phi_\mu(x, z) = \tfrac{1}{2}\|Ax - b\|^2 + \sum_{i=1}^{n}\Big(x_i z_i - \mu \ \ln(x_i z_i)\Big). \end{aligned} \tag{7}$$

It is apparent from the way the problem is formulated, that the variables x and z are positive and therefore the function Φ_μ is well defined.

Property 4.2. *For fixed $\mu > 0$, $n\mu(1 - ln(\mu))$ is the global minimum of the function Φ_μ and is attained at each point on the quasicentral path. In other words,*

$$\min \ \Phi_\mu(x, z) = n\mu(1 - \ln(\mu)) = \Phi_\mu(x^*_\mu, z^*_\mu)$$

for each (x^*_μ, z^*_μ) on the quasicentral path.

Proof.

It is easy to verify that $\Phi_\mu(w) = w - u \ln w, w > 0$ attains its global minimum at $w = \mu$. Therefore $\sum_{i=1}^{n}(x_i z_i - \mu \ln(x_i z_i))$ attains its global minimum, $n\mu(1 - \ln \mu)$, at every point (x, z) on the quasicentral path. It follows that

$$\Phi_\mu(x, z) \geq n\mu(1 - \ln(\mu)).$$

The conclusion follows since we have $\Phi_\mu(x^*_\mu, z^*_\mu) = n\mu(1 - \ln(\mu))$ at each point (x^*_μ, z^*_μ) on the quasicentral path.

□

Property 4.3. *For fixed $\mu > 0$, the Newton direction $\Delta\tilde{v} = (\Delta x, \Delta z)$ is a descent direction for Φ_μ at each positive point $\tilde{v} = (x, z)$ not on the quasicentral path, i.e.,*

$$\nabla\Phi_\mu(\tilde{v})^T\Delta\tilde{v} < 0.$$

Proof. The components of the gradient of Φ_μ with respect to x and z are

$$\nabla_x\Phi_\mu(x, z) = A^T(Ax - b) + z - \mu x^{-1} \quad \text{and} \quad \nabla_z\Phi_\mu(x, z) = x - \mu z^{-1}.$$

The directional derivative of Φ_μ in the direction $\Delta\tilde{v} = (\Delta x, \Delta z)$ at $\tilde{v} = (x, z)$ is given by

$$\nabla\Phi_\mu(x, z)^T \begin{pmatrix} \Delta x \\ \Delta z \end{pmatrix} = \nabla_x\Phi_\mu(x, z)^T\Delta x + \nabla_z\Phi_\mu(x, z)^T\Delta z.$$

$$= (Ax - b)^T A\Delta x + (z - \mu x^{-1})^T\Delta x + (z - \mu z^{-1})^T\Delta z.$$

If we set $W = (XZ)^{-1/2}$, and by using the first and third block of equations of (6), we obtain

$$\nabla\Phi_\mu(x, z)^T \begin{pmatrix} \Delta x \\ \Delta z \end{pmatrix} = -\Big(\|Ax - b\|^2 + \|W(XZe - \mu e)\|^2\Big) < 0. \tag{8}$$

This inequality establishes the theorem. □

Sufficient Decrease. Since Φ_μ is a continuously differentiable function by Proposition 4.2 bounded from below, and by Proposition 4.3 since the Newton direction $\Delta\tilde{v} = (\Delta x, \Delta z)$ is a descent direction for Φ_μ, then it is known from [5] that for any fraction $\beta \in (0, 1)$, there exists an $\alpha^* > 0$ such that the following rate of decrease

$$\Phi_\mu(\tilde{v} + \alpha\Delta\tilde{v}) \leq \Phi_\mu(\tilde{v}) + \beta\alpha\nabla\Phi_\mu(\tilde{v})^T\Delta\tilde{v} \tag{9}$$

holds for any $\alpha \in (0, \alpha^*]$.

A continuation, we prove that the merit function Φ_μ plays a key role in preventing that the sequence $\{X^k Z^k e, k \in \mathbb{N}\}$, generated by the Algorithm 1, goes to zero or infinity for any fixed $\mu > 0$.

Property 4.4. *For fixed* $\mu > 0$*, the sequence* $\{X^k Z^k e, k \in \mathbb{N}\}$ *is bounded and bounded away from zero.*

Proof. From inequality (9), we know that the sequence $\{\Phi_\mu(x^k, z^k), k \in \mathbb{N}\}$ is non-increasing, and since $\Phi_\mu(x^k, z^k)$ is bounded below by $n\mu(1 - \ln \mu)$, then

$$n\mu(1 - \ln \mu) \leq \Phi_\mu(x^k, z^k) \leq \Phi_\mu(x^o, z^o).$$

If $x_j^k z_j^k \to 0$ or ∞ then $x_j^k z_j^k - \mu \ln(x_j^k z_j^k) \to \infty$. This contradicts the above inequality. Thus, there exists a positive constant C such that for every $k = 1, 2, \ldots,$

$$\frac{1}{C} \leq x^k z^k \leq C. \tag{10}$$

This concludes the proof. $\square$

5 Proximity to the Quasicentral Path

It is important to observe that the absolute value of the directional derivative of Φ_μ in any Newton direction $\Delta\tilde{v}$ can be interpreted as a weighted deviation from the quasicentral path. Therefore we use this value as a measure of proximity to the quasicentral path. We formalize this idea with the following definition.

Definition 5.1. *We say that a positive point* (x, z) *is sufficiently close to the quasicentral path if*

$$\mathcal{N}_W(\gamma\mu) = \Big\{(x, z) \in \mathrm{I\!R}^{n+n} : \|Ax - b\|^2 + \|W(XZe - \mu e)\|^2 \leq \gamma\mu\Big\}$$

where $W = (XZ)^{-1/2}$*, and* $\gamma \in (0, 1)$*.*

In particular if $W = I$, the set defined above becomes

$$\mathcal{N}_2(\gamma\mu) = \Big\{(x, z) \in \mathrm{I\!R}^{n+n} : \|Ax - b\|^2 + \|(XZe - \mu e)\|^2 \leq \gamma\mu\Big\},$$

and this set can be interpreted as a deviation from the quasicentral path measured in the 2-norm.

In order to facilitate the comparison between $\mathcal{N}_W(\gamma\mu)$ and $\mathcal{N}_2(\gamma\mu)$, we introduce the following definitions.

Definition 5.2. *We say that a positive point* (x, z) *is far away from the solution set if* $x_i z_i > 1$, *for* $i = 1, \ldots, n$.

Definition 5.3. *We say that a positive point* (x, z) *is close enough to the solution set if* $x_i z_i < 1$, *for* $i = 1, \ldots, n$.

Now we express the relationship between $\mathcal{N}_W(\gamma\mu)$ and $\mathcal{N}_2(\gamma\mu)$.

Property 5.4. *For* $\mu > 0$, *if* $0 < x_i z_i \leq 1$, $i = 1, \ldots, n$, *then* $\mathcal{N}_W(\gamma\mu) \subseteq \mathcal{N}_2(\gamma\mu)$.

Proof. Since $0 < x_i z_i \leq 1$, then $0 < (x_i z_i - \mu)2 \leq (x_i z_i - \mu)2/(x_i z_i)$. Therefore

$$0 < \|XZe - \mu e\| \leq \|W(XZe - \mu e)\|.$$

The proof follows directly from the above inequality. □

Property 5.5. *For* $\mu > 0$, *if* $x_i z_i > 1$, $i = 1, \ldots, n$, *then* $\mathcal{N}_2(\gamma\mu) \subseteq \mathcal{N}_W(\gamma\mu)$.

Proof. Since $x_i z_i > 1$, then $(x_i z_i - \mu)2/(x_i z_i) \leq (x_i z_i - \mu)2$. Therefore

$$0 < \|W(XZe - \mu e)\| \leq \|XZe - \mu e\|.$$

The proof follows as that of Property 5.1. □

From Properties 5.4 and 5.4 it is readily concluded that far away from the solution set, weighted neighborhoods are contained in the 2-norm neighborhoods, whereas near the solution set, the 2-norm neighborhoods are contained in the weighted neighborhoods. Therefore, near to a solution, the use of weighted neighborhoods may allow larger step lengths. This is the reason that we are proposing weighted neighborhoods as a measure of closeness to the quasicentral path.

6 Global Convergence Theory

The global convergence theory of Algorithm 1 is proved by using as assumption that the sequence $\{(x^k, z^k)\}$ is bounded. We prove that for a fixed $\mu > 0$, any limit point of the sequence generated by the inner loop algorithm is on the quasicentral path. Moreover, any limit point of the sequence generated by the outer loop algorithm converges to a solution of the primal and dual problems.

Assumption 6.1. *The sequence* $\{(x^k, z^k), k \in \mathbb{N}\}$ *generated by the Algorithm 1 is bounded.*

Corollary 6.1. *Let* $\{(x^k, z^k), k \in \mathbb{N}\}$ *be the sequence generated by the inner loop of Algorithm 1. Then there exists a constant* $M > 0$ *such that for* $k \in \mathbb{N}$, *the inequalities*

$$\frac{1}{M} \leq \|x^k\| \leq M \ \ \textit{and} \ \ \frac{1}{M} \leq \|z^k\| \leq M$$

hold. In other words, each sequence is bounded and bounded away from zero.

Proof. The proof follows immediately from the Assumption 6.1 and inequality (10). □

Theorem 6.1. *For fixed* $\mu > 0$, *any limit point* (x^*_μ, z^*_μ) *of the sequence* $\{(x^k, z^k), k \in \mathbb{N}\}$ *generated by the inner loop of Algorithm 1, with the stopping criterion deactivated, is on the quasicentral path, i.e.,*

$$Ax^*_\mu - b = 0 \ \ \textit{and} \ \ X^*_\mu Z^*_\mu e = \mu e, \tag{11}$$

and

$$x^*_\mu > 0, z^*_\mu > 0.$$

Proof. Let (x^*_μ, z^*_μ) be a limit point of the sequence $\{\{(x^k, z^k), k \in \mathbb{N}\}$, and choose a subsequence, still denoted by $\{(x^k, z^k), k \in \mathbb{N}\}$, converging to (x^*_μ, z^*_μ). From the Corollary 6.1, we know that the sequences $(x^k)_k$ and $(z^k)_k$ are bounded away from zero, which implies

$$x^*_\mu > 0 \quad \text{and} \quad z^*_\mu > 0.$$

We will now show that the point (x^*_μ, z^*_μ) satisfies the conditions (11).

We start by observing that the sequence $\{(\Delta x^k, \Delta z^k), k \in \mathbb{N}\}$ is bounded. To see this, we observe that since (x^k, y^k, z^k) is an interior point and A is full rank, then $F^{'}_\mu$ is a nonsingular matrix. Thus for each $k = 1, 2, \dots,$ we have

$$(\Delta x^k, \Delta y^k, \Delta z^k) = -\left(F^{'}_\mu(x^k, y^k, z^k)\right)^{-1} F_\mu(x^k, y^k, z^k).$$

Since F_μ and $(F^{'}_\mu)^{-1}$ are continuous functions on $\mathrm{I\!R}^{n+m+n}$ and the set

$$E = \left\{(x^k, y^k, z^k), k \in \mathbb{N}\} \cup \{(x^*_\mu, y^*_\mu, z^*_\mu)\right\}$$

is compact, then the sequence $\{(\Delta x^k, \Delta y^k, \Delta z^k), k \in \mathbb{N}\}$ is bounded, being the image of a compact set under a continuous function. Thus, its projection on the $x - z$ plane, namely the set $\{(\Delta x^k, \Delta z^k), k \in \mathbb{N}\}$, is bounded.

Next, we claim that the step length sequence $\{\alpha_k, k \in \mathbb{N}\}$ generated by Algorithm 1 in Step 4 is bounded away from zero. If this were not so, then there must be a subindex $1 \leq j \leq n$ for which

$$\lim_{k\to\infty} \frac{x^k_j}{\Delta x^k_j} = 0.$$

Since the sequence x^k_j is bounded away from zero, then the previous equation implies that

$$\lim_{k\to\infty} \Delta x^k_j = \infty.$$

This contradicts the boundedness of the sequence $\{(\Delta x^k, \Delta z^k), k \in \mathbb{N}\}$.

Now, since Φ_μ is a continuously differentiable function on $\mathrm{I\!R}^{n+n}$, bounded below in the class of all positive points (x, z), and as we are using the following iterative scheme

$$(x^{k+1}, z^{k+1}) = (x^k, z^k) + \alpha_k(\Delta x^k, \Delta z^k) > 0,$$

then we have

$$\lim_{k\to\infty} \frac{\nabla \Phi_\mu(x^k, z^k)^T \begin{pmatrix} \Delta x^k \\ \Delta z^k \end{pmatrix}}{\|(\Delta x^k, \Delta z^k)\|} = 0.$$

Since $\{(\Delta x^k, \Delta z^k), k \in \mathbb{N}\}$ is bounded, we have

$$\lim_{k\to\infty} \nabla\Phi_\mu(x^k, z^k)^T \left(\begin{array}{c} \Delta x^k \\ \\ \Delta z^k \end{array} \right) = -\lim_{k\to\infty} \left(\|Ax^k - b\|^2 + \|W^k(X^kZ^ke - \mu e)\|^2 \right) = 0.$$

From (10), we have $0 < C \leq (X^kZ^k)^{-1/2}$ for $k \in \mathbb{N}$. Therefore

$$C\|X^kZ^ke - \mu e\| \leq \|(X^kZ^k)^{-1/2}(X^kZ^ke - \mu e)\| \quad \text{for} \ \ k \in \mathbb{N}.$$

By continuity of the norm and the previous result, we conclude that

$$Ax^* = b \quad \text{and} \quad X^*_\mu Z^*_\mu e = \mu e. \qquad \Box$$

Comment 6.1. *Following the quasicentral path exactly, for every element of a sequence* $\{\mu^k, k \in \mathbb{N}\}$ *which converges to zero, carries a high computational cost. To avoid such an expensive calculation, we activate a stopping criterion for obtaining a suitable approximation to the quasicentral path.*

We adopt the following convention. Consider a sequence $\{\mu^k, k \in \mathbb{N}\}$ such that $\mu^k \to 0$. Now for a fixed k, let $\{(x^{k^i}, z^{k^i}), i \in \mathbb{N}\}$ be the sequence generated by the inner loop of Algorithm 1. For a given $\mu^k > 0$, the stopping criterion of the inner loop algorithm is activated when a positive point (x^{k^i}, z^{k^i}) satisfies the following inequality

$$\|Ax^{k^i} - b\|^2 + \|W^{k^i}(X^{k^i}Z^{k^i}e - \mu^k e\|^2 \leq \gamma\mu^k, \ \ \gamma \in (0,1) \tag{12}$$

We let

$$l_k = \inf\{i : (x^{k^i}, z^{k^i}) \in \mathcal{N}_W(\gamma\mu^k)\}$$

and rewrite

$$(x^{k^{l_k}}, z^{k^{l_k}}) \ \text{ as } \ (x^k, z^k).$$

In other words, for a given μ^k the positive point (x^k, z^k) is the first element generated by the inner loop of Algorithm 1, that satisfies the inequality (12).

The sequence $\{\mu^k, k \in \mathbb{N}\}$ is chosen in the following way: Provided that the point (x, z)

is not a solution of the problem, once it lies inside the weighted neighborhood (i.e., (x, z) satisfies the inequality in Step 5), the parameter μ is updated as

$$\mu_{k+1} = \sigma_k \Big\{ \|Ax^k - b\| + \|W^k(X^k Z^k e - \mu^k e)\| \Big\}, \tag{13}$$

where $0 < \sigma_k < 1$. It is then easily verified that the sequence $\{\mu^k, k \in \mathbb{N}\}$ is superlinearly convergent to zero. We now prove that any limit point of the sequence generated by the outer loop of Algorithm 1 when the perturbed parameter μ goes to zero is a solution of the primal and dual problems.

Theorem 6.2. *Consider a limit point (x^*, z^*) of the sequence $\{(x^k, z^k),\ k \in \mathbb{N}\}$ generated by the outer loop of Algorithm 1 as the perturbed parameter μ^k goes to zero. Then (x^*, z^*) satisfies the primal and complementarity conditions, that is*

$$Ax^* = b \ \ \textit{and} \ \ X^* Z^* e = 0.$$

Furthermore, there exists a sequence $\{y^k, k \in \mathbb{N}\}$ implicity associated with $\{(x^k, z^k), k \in \mathbb{N}\}$ generated by Algorithm 1 such that

$$y^k \to y^* \ \ \textit{as} \ \ k \to \infty$$

where

$$A^T y^* + z^* - c = 0.$$

Proof. Since (x^*, z^*) is a limit point of the sequence $\{(x^k, z^k), k \in \mathbb{N}\}$ generated by the Algorithm 1, there exists a convergent subsequence $\{(x^k, z^k);\ k \in \mathbb{N}\}$ such that $(x^k, z^k) \to (x^*, z^*)$, where $(x^k, z^k) \in \mathcal{N}_W(\gamma\mu^k)$ and $\mu^k \to 0$.
Now, since $(x^k, z^k) \in \mathcal{N}_W(\gamma\mu^k)$, we have the inequality

$$\|Ax^k - b\|^2 + \|(X^k Z^k)^{-1/2}(X^k Z^k e - \mu^k e)\|^2 \leq \gamma\mu^k.$$

Taking the limit, we obtain

$$\lim_{k\to\infty} \|Ax^k - b\|^2 + \lim_{k\to\infty} \|(X^k Z^k)^{-1/2}(X^k Z^k e - \mu^k e)\|^2 \leq 0. \tag{14}$$

By continuity of the norm

$$\lim_{k\to\infty} \|Ax^k - b\|^2 = \|Ax^* - b\|^2 = 0,$$

which implies

$$Ax^* = b.$$

Owing to Proposition 6.1, we have

$$\|X^k Z^k\| \leq C, k = 1, 2, \ldots$$

for a positive constant C. It follows that

$$\frac{1}{\sqrt{C}} \leq \|X^k Z^k\|^{-1/2}.$$

Thus,

$$\frac{1}{\sqrt{C}}\|X^k Z^k e - \mu^k e\|^2 \leq \|(X^k Z^k)^{-1/2}(X^k Z^k - \mu^k e)\|^2.$$

Taking limits of both sides and exploiting the continuity of the norm, (19) yields

$$X^* Z^* e = 0.$$

This finishes the proof of the first part of the theorem.

As for the second part, consider now the implicitly generated sequence $\{y^k, k \in \mathbb{N}\}$. We know that $A^T y^k + z^k - c \to 0$ since $Ax^k - b \to 0$ as $k \to \infty$. The expression

$$A^T y^k = (A^T y^k + z^k - c) + c - z^k$$

together with the fact that $\lim\limits_{k\to\infty} z^k = z^*$ yields that $\lim\limits_{k\to\infty} A^T y^k = c - z^*$.

Since the $\mathcal{R}(A^T) = \{A^T y \text{ for any } y \in \mathrm{I\!R}^m\}$ is a closed subspace of $\mathrm{I\!R}^n$, we conclude immediately that

$$\lim_{k\to\infty} A^T y^k \in \mathcal{R}(A^T).$$

It follows that there exists $y^* \in \mathrm{I\!R}^m$ such that

$$\lim_{k\to\infty} A^T y^k = A^T y^* = c - z^*.$$

Since the columns of A^T are linearly independent, then y^* is the unique vector in $\mathrm{I\!R}^m$ that satisfies the above condition. This concludes the proof. $\square$

7 Numerical Experimentation

In this section we show how Algorithm 1 presented in Section 3 performs numerically in obtaining a solution for a set of test problems. It is important to state that our current goal is not to compare the numerical behavior with other algorithms, but to show that using the quasicentral path as a central region it is enough for guiding the iterates towards a solution of the problem. Now, Algorithm 1 was written in MATLAB version 6a, and the implementation was done on a Sun Ultra 10 machine running the Solaris system. The numerical experiments were performed on the set of NETLIB test problems. In Tables 1-4, we summarized the numerical results obtained by the Algorithm 1 where the first four columns contains the problem number, problem name and dimensions of the problem, respectively. The next two columns state the number of linear systems solved by Algorithm 1 and its corresponding cpu time in seconds. Finally, the last three columns denote the primal objective, and the norms of the primal and dual conditions at the initial point. The primal conditions are denoted by $e_p^0 = \|Ax_0 - b\|$ and the dual conditions are given by be $e_d^0 = \|z - c\|$ where the initial value of the variable y_0 is set to zero.

Now, this implementation of the algorithm entails the selection of an initial interior point (x_0, y_0, z_0) satisfying the inequality

$$\|z_0 - c\| \leq \|Ax_0 - b\|. \tag{15}$$

The initial point is chosen by following a procedure widely used in the literature: we take $y_0 = 0$, then pick x_0 and z_0. If the point $(x_0, 0, z_0)$ satisfies the condition (15), we let this be our initial point.

Otherwise, we solve for ξ the following inequality

$$\|z_0 - c\| \leq \|A(\xi x_0) - b\|, \tag{16}$$

which is equivalent to

$$\xi 2\|Ax_0\|^2 - 2\xi\langle Ax_0, b\rangle + \|b\|^2 - \|z_0 - c\|^2 \geq 0. \tag{17}$$

It is easy to verify that in the present situation, the equation

$$\xi 2\|Ax_0\|^2 - 2\xi\langle Ax_0, b\rangle + \|b\|^2 - \|z_0 - c\|^2 = 0 \tag{18}$$

has two real distinct roots $\xi_1 < 1 < \xi_2$.

If $\xi_1 \leq 0$, the inequality (17) holds as long as $\xi \leq \xi_2$, then we set $\xi = \xi_2$. Else, we might take

$$\xi \in (0, \xi_1] \cup [\xi_2, \infty)$$

and set $(\xi x_0, 0, z_0)$ as our initial point where $\xi = 10 * \max\{\xi_1, \xi_2\}$. In the previous conclusions, we assume that $A * x_0 \neq 0$. The parameters τ and γ are set to 0.99995 and .5, respectively.

8 Conclusions

In this work we have presented an infeasible primal-dual interior-point method for solving linear programs, a global convergence theory, and a numerical experimentation of the strategy. We show that the use of the quasicentral path, while being a less restrictive notion than the central path, it is sufficiently strong to guide the iterates towards a solution of the problem. Moreover, our methodology includes a new merit function and weighted proximity measures. The numerical results supports the proposed globalization strategy for solving linear programming problems. Future works includes more numerical experimentation and comparisons with other strategies for solving large scale linear programs.

References

[1] ArgáeZ, M. 1997. Exact and inexact Newton linesearch interior–point methods for nonlinear programming. Ph.D Thesis, Department of Computational and Applied Mathematics, Rice University.

[2] Argáez, M. and Tapia, R.A. 2002. On the global convergence of a modified augmented Lagrangian linesearch interior–point Newton method for nonlinear programming. J. Optimi. Theory Appl. Vol.114(1), 1–25.

[3] Argáez, M., Tapia, R.A., and Velźquez, L. 2002. Numerical comparisons of path–following strategies for a primal–dual interior–point method for nonlinear programming. J. of Optimi. Theory Appl. Vol.114(2), 255–272.

[4] Bixby, R. 2002. Solving Real-World Linear Programs: A Decade and More of Progress. Operations Research Vol. 50(1), 3–15.

[5] Dennis, J. E. Jr. and Schnabel, R. B. 1996 *Numerical methods for unconstrained optimization and nonlinear equations.* SIAM, Philadelphia, PA.

[6] EL-Bakry, A. S., Tapia, R. A., Tsuchiya T., and Zhang, Y. 1996. On the formulation of the primal-dual interior point method for nonlinear programming. J. of Optimi. Theory Appl. Vol. 89, 507-541.

[7] Karmarkar, N. 1994. *A new polynomial-time algorithm for linear programming.* Combinatorica 4, 373-395.

[8] Kojima, M., Mizuno, S., and Yoshise, A. 1989. A primal-dual interior point method for linear programming. in Progress in Mathematical Programming: Interior-Point and Related Methods, N. Megiddo, ed., Springer-Verlag, New York-Berlin, Ch.2, 29–47.

[9] Lustig, I.J., Marsten, R., Shanno, D.F. 1994. interior point methods for linear programming: computational state of the art. ORSA J. Comput., Vol. 6(1), 1–14.

[10] Ortega, J. M. and Rheinboldt, W. C. 1970. Iterative solution of nonlinear equations in several variables. Academic Press, New York, NY.

[11] Wright, S.J. 1997. Primal-dual interior-point methods. SIAM, Philadelphia, PA.

[12] Zhang, Y. 1995. *User's guide to LIPSOL.* Technical Report TR95-19, Dept. of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, MD 21228.

List of Tables

Table 1: Numerical Results

Name	m	n	Iter	CPU	Residual	$\|e_d^0\|$	$\|e_p^0\|$
25fv47	798	1854	24	5.18	5.94e-13	2.78e+03	2.74e+04
80bau3b	2235	11516	40	21.91	8.10e-09	8.72e+04	6.07e+05
adlittle	55	137	19	0.57	1.57e-12	1.93e+04	3.92e+04
afiro	27	51	8	0.29	4.10e-10	3.07e+01	2.94e+03
agg	488	615	21	1.99	3.93e-12	6.73e+03	3.78e+07
agg2	516	758	19	2.54	6.13e-09	9.45e+03	7.41e+06
agg3	516	758	18	2.39	4.25e-11	9.45e+03	7.42e+06
bandm	269	436	18	0.90	.57e-09	2.37e+02	2.56e+03
beaconfd	148	270	14	0.98	1.72e-10	9.35e+02	3.97e+05
blend	74	114	13	0.51	1.23e-12	4.00e+01	2.06e+03
bnl1	632	1576	25	2.60	2.41e-12	8.41e+02	1.83e+04
bnl2	2268	4430	32	14.85	1.12e-11	3.13e+03	7.54e+04
boeing1	347	722	22	1.85	4.31e-09	4.47e+01	2.18e+04
boeing2	140	279	18	0.84	1.74e-12	2.95e+01	5.48e+05
bore3d	199	300	18	1.19	3.45e-11	2.62e+03	7.01e+04
brandy	149	259	16	0.70	1.24e-11	9.67e+00	3.57e+03
capri	267	476	20	1.38	9.23e-12	2.61e+01	1.69e+04
cycle	1801	3305	29	15.91	3.61e-09	6.96e+01	2.24e+04
czprob	737	3141	36	4.95	5.56e-11	5.66e+04	1.29e+06
d2q06c	2171	5831	31	32.93	1.83e-12	1.63e+04	1.97e+05
d6cube	404	6184	31	16.28	8.38e-11	2.63e+03	3.91e+06
degen2	444	757	16	1.72	3.94e-11	2.85e+03	1.40e+04
degen3	1503	2604	27	25.31	7.48e-09	2.40e+03	4.75e+04
dfl001	6071	12230	48	1900.29	1.71e-08	1.11e+05	1.18e+05
e226	220	469	21	0.94	1.43e-12	3.71e+02	3.47e+03

Table 2: Numerical Results (Cont.)

Name	m	n	Iter	CPU	Residual	$\|e_d^0\|$	$\|e_p^0\|$
etamacro	357	692	27	2.11	7.40e-10	2.93e+03	3.93e+03
fffff	800	501	24	3.04	1.79e-09	3.81e+01	3.16e+07
finnis	492	1014	30	1.93	1.51e-09	3.13e+04	2.96e+05
fit1d	24	1049	19	2.95	5.31e-09	7.72e+02	1.26e+04
fit1p	627	1677	16	16.19	3.30e-10	8.11e+02	1.67e+03
fit2d	25	10524	43	53.88	5.79e-10	3.20e+03	1.34e+05
fit2p	3000	13525	14	178.19	1.08e-10	3.26e+03	6.54e+05
forplan	135	463	22	1.26	4.80e-11	6.93e+01	2.14e+06
ganges	1137	1534	19	3.24	2.10e-11	7.48e+01	2.89e+06
giffpin	600	1144	19	1.57	3.14e-10	2.98e+04	1.13e+09
greenbea	2318	5424	100	142.59	5.85e-06	4.04e+04	8.23e+05
greenbeb	2317	5415	37	18.73	1.26e-09	4.05e+04	7.16e+05
grow15	300	645	15	1.76	7.05e-10	8.71e+01	1.39e+03
grow22	440	946	16	2.86	1.15e-09	1.30e+02	1.67e+03
grow7	140	301	15	1.04	5.88e-09	3.98e+01	9.64e+02
israel	174	316	23	1.99	4.63e-12	1.78e+04	1.21e+06
kb2	43	68	14	0.79	2.37e-10	8.46e+01	2.49e+05
lotfi	151	364	18	0.64	6.25e-10	1.99e+01	4.15e+04
maros	835	1921	31	4.90	2.84e-10	6.62e+03	7.56e+04
maros-r7	3136	9408	14	181.33	1.08e-10	3.26e+03	6.54e+05
modszk1	686	1622	24	2.15	1.19e-09	4.03e+04	5.24e+06
nesm	654	2922	34	8.09	4.46e-09	3.64e+04	7.64e+05
perold	625	1530	57	7.93	4.30e-12	1.30e+00	2.84e+05
pilot	1441	4657	32	52.85	2.51e-09	5.16e+00	6.24e+04
pilot4	402	1173	51	6.35	1.34e-09	8.42e-01	1.20e+04
pilot87	2030	6460	36*	170.58	5.31e-08	7.20e+02	4.23e+05
pilotja	924	2044	41	13.23	4.04e-10	1.15e+00	2.60e+05

Table 3: Numerical Results (Cont.)

Name	m	n	Iter	CPU	Residual	$\|e_d^0\|$	$\|e_p^0\|$
pilotnov	951	2242	22	7.46	7.85e-12	6.35e+01	4.48e+05
pilotwe	722	2930	35	5.55	3.77e-09	1.38e+05	9.48e+05
recipe	85	177	10	0.61	6.58e-12	2.35e+01	2.75e+05
sc105	105	163	10	0.41	1.62e-10	5.77e+00	1.63e+03
sc205	205	317	11	0.63	7.03e-13	7.82e+00	2.38e+03
sc50a	49	77	8	0.42	1.45e-10	4.21e+00	1.13e+03
sc50b	48	76	7	0.62	3.51e-11	4.19e+00	1.47e+03
scagr25	471	671	15	0.93	3.26e-09	2.76e+04	1.89e+05
scagr7	129	185	15	0.98	3.71e-11	1.43e+04	1.03e+05
scfxm1	322	592	18	1.06	5.53e-13	3.60e+02	5.06e+04
scfxm2	644	1184	20	1.87	3.16e-11	7.20e+02	7.19e+04
scfxm3	966	1776	20	2.55	1.16e-11	1.08e+03	8.81e+04
scorpion	375	453	15	0.75	1.49e-10	7.43e+03	1.49e+04
scrs8	485	1270	26	2.03	4.04e-11	3.56e+04	3.83e+06
scsd1	77	760	10	0.90	1.17e-12	0.00e+00	1.00e+00
scsd6	147	1350	15	1.82	8.34e-09	0.00e+00	1.98e+00
scsd8	397	2750	11	2.50	7.59e-12	0.00e+00	1.28e+01
sctap1	300	660	17	0.64	5.44e-13	4.86e+03	7.07e+04
sctap2	1090	2500	16	2.30	1.50e-10	1.92e+04	1.45e+05
sctap3	1480	3340	19	3.40	1.12e-11	2.32e+04	1.53e+05
seba	515	1036	23	6.49	2.81e-10	1.20e+04	2.58e+05

Table 4: Numerical Results (Cont.)

Name	m	n	Iter	CPU	Residual	$\|e_d^0\|$	$\|e_p^0\|$
share1b	112	248	21	0.74	1.53e-11	6.79e+02	7.70e+05
share2b	96	162	14	0.41	2.03e-12	4.48e+01	3.54e+03
shell	496	1487	20	1.70	2.01e-12	3.70e+04	8.16e+06
ship04l	356	2162	12	1.53	1.67e-09	4.65e+04	5.20e+04
ship04s	268	1414	13	1.41	3.55e-09	3.76e+04	7.52e+04
ship08l	688	4339	15	3.86	4.12e-12	6.59e+04	7.42e+04
ship08s	416	2171	14	1.98	8.40e-11	4.66e+04	9.32e+04
ship12l	838	5329	16	4.00	3.62e-12	7.30e+04	7.36e+04
ship12s	466	2293	18	2.42	2.55e-11	4.79e+04	9.57e+04
siera	1222	2715	19	5.21	1.28e-11	2.64e+04	3.83e+05
stair	356	538	15	1.61	6.37e-13	1.00e+01	2.85e+03
standata	359	1258	18	1.54	6.73e-09	2.50e+03	2.69e+06
standgub	361	1366	18	1.18	3.66e-10	2.62e+03	2.69e+06
standmps	467	1258	26	1.62	3.64e-12	2.50e+03	2.48e+06
stocfor1	109	157	14	0.55	1.96e-10	5.18e+03	5.32e+05
stocfor2	2157	3045	21	4.66	2.21e-10	4.32e+03	2.38e+06
stocfor3	16675	23541	34	46.55	2.35e-09	6.52e+03	6.63e+06
truss	1000	8806	20	9.07	5.90e-10	6.83e+04	1.37e+05
tuff	292	617	15	1.47	9.17e-10	8.78e-03	2.85e+04
vtp.base	194	325	45	1.93	1.55e-11	1.55e+01	1.22e+05
wood1p	244	2595	21	14.41	329e-10	2.16e+01	7.65e+04
woodw	1098	8418	28	13.68	1.74e-11	7.80e+01	1.23e+05