LOCATING THE LEAST 2-NORM SOLUTION OF LINEAR PROGRAMS VIA A PATH-FOLLOWING METHOD *

YUN-BIN ZHAO † AND DUAN LI‡

Abstract. A linear program has a unique least 2-norm solution provided that the linear program has a solution. To locate this solution, most of the existing methods were devised to solve certain equivalent perturbed quadratic programs or unconstrained minimization problems. Different from these traditional methods, we provide in this paper a new theory and an effective numerical method to seek the least 2-norm solution of a linear program. The essence of this method is a (interior-point-like) path-following algorithm that traces a newly introduced regularized central path which is fairly different from the central path used in interior-point methods. One distinguishing feature of the method is that it imposes no assumption on the problem. The iterates generated by this algorithm converge to the least 2-norm solution whenever the linear program is solvable; otherwise, the iterates converge to a point which gives a minimal KKT residual when the linear program is unsolvable.

Key words. Linear programming, path-following algorithm, regularized central path, least 2-norm solution.

AMS subject classifications. 90C05, 90C33, 90C51, 65K05

1. Introduction. Consider the linear program:

$$\min\{c^T x : \ Ax \geq b, \ x \geq 0\}, \tag{1.1}$$

where $A \in R^{m\times n}$, $c \in R^n$ and $b \in R^m$. The dual problem for the above linear program can be written as

$$\max\{b^T y : \ A^T y \leq c, \ y \geq 0\}. \tag{1.2}$$

Let S_P^* and S_D^* denote the optimal solution sets (possibly empty) of the problems (1.1) and (1.2), respectively. If a linear program has an optimal solution, it is said to be solvable; otherwise, it is unsolvable. According to linear programming theory (see for instance, Theorem 1.13 in [34]), the primal (1.1) and the dual (1.2) have optimal solutions if and only if both problems have feasible solutions. If one of the problem (1.1) or (1.2) has no feasible solution, then the other one is either unbounded or has no feasible solution, and if one of the problem (1.1) or (1.2) is unbounded, then the other one has no feasible solution. Therefore, we may say that the primal problem is solvable if and only if the dual is solvable. Equivalently, the primal is unsolvable if and only if the dual is unsolvable.

Throughout this paper, we denote by $\|\cdot\|_\infty$ the ∞-norm of a vector, and $\|\cdot\|_2$ the 2-norm, i.e., Euclidean norm. The purpose of this paper is to give a new method to find the least 2-norm solutions of both primal and dual linear programs, i.e., to find $x^* \in S_P^*$ and $y^* \in S_D^*$ such that

$$\|x^*\|_2 \leq \|x\|_2, \ \|y^*\|_2 \leq \|y\|_2 \ \text{ for all } x \in S_P^* \text{ and } y \in S_D^*.$$

*This work was partially supported by Grant CUHK4392/99E, Research Grants Council, Hong Kong.

† Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, New Territory, Hong Kong, and Institute of Applied Mathematics, AMSS, Chinese Academy of Sciences, Beijing, China (Email: ybzhao@se.cuhk.edu.hk).

‡ Corresponding author. Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong. Shatin, New Territory, Hong Kong (Fax: (852) 26035505; Tel: (852) 26098323; Email: dli@se.cuhk.edu.hk).

Given certain norm, the problem of finding the least-norm solution to some optimization problems or other applied mathematical problems have been studied by many authors such as Tikhonov and Arsenin [30], Tucker [28, 29], Parsons and Tucker [21], and Wolfe [32, 33]. In particular, many authors have studied the theoretical property of the least 2-norm solution of a linear program, and have tried to design numerical methods to compute this solution. See for instance, Tikhonov and Arsenin [30], Mangasarian [13, 14, 15, 16], Mangasarian and Meyer [19], Mangasarian and De Leone [18], Lucidi [11, 12], Skarin [25], Kiwiel [6, 7], Smith and Wolkowicz [24], and Kanzow, Qi and Qi [5]. Note that the least 2-norm solution of a linear program could be a vertex of the feasible set, and could be also a relative interior point of the optimal faces. Thus, in general case, both simplex methods and interior-point methods (see [23, 34]) may not find the least 2-norm solution of a linear program.

The first method for the least-norm solution of a linear program was the canonical Tikhonov regularization method [30]. The basic idea of this method is to solve successively the following quadratic problem in x:

$$\min\{c^T x + \mu\|x\|_2^2 : \ Ax \geq b, \ x \geq 0\}, \tag{1.3}$$

where μ is a positive parameter. For each $\mu > 0$, denote by $x(\mu)$ the solution to the above quadratic program. Tikhonov (see [30]) showed that $x(\mu)$ converges, as $\mu \to 0$, to the least 2-norm solution of (1.1). Later, Mangasarian and Meyer [19] showed that there exists a $\bar{\mu} > 0$ such that for any $\mu \in (0, \bar{\mu}]$, the perturbed quadratic program (1.3) becomes an exact problem, i.e., for any $\mu \in (0, \bar{\mu}]$ the solution $x(\mu) = \bar{x}$, where $\bar{x}$ is the least 2-norm solution of (1.1). Based on this observation, Mangasarian [14, 16] used successive overrelaxation (SOR) methods to solve the dual problem of (1.3). As pointed out by Lucidi [11], the main advantage of SOR algorithms is that they preserve the sparsity structure of the problem, and thus can tackle large scale problems. However, the main difficulty encountered by this method appears to be the difficulty of knowing such a threshold value of $\bar{\mu}$. Thus, in general, it is not sure if a value of μ is small enough such that the solution $x(\mu)$ of (1.3) is the least 2-norm solution of (1.1). Even the condition $x(\mu^{k+1}) = x(\mu^k)$ with $\mu^{k+1} < \mu^k$ does not imply that $x(\mu^{k+1})$ is the least 2-norm solution of (1.1).

There are two classes of ways to circumvent this difficulty. The first class of approaches, including Lucidi [11, 12] and Kiwiel [6], attempts to establish an effective computational criterion to check whether the current perturbed quadratic program is exact. However, Lucidi's methods [11, 12] require that the linear program (1.1) be nondegenerate (the gradients of active constrains at the least 2-norm solution $\bar{x}$ are linearly independent), whereas Kiwiel's method [6] solves the perturbed quadratic program by finite active-set methods (see for example, Best [1] and Kiwiel [8]) which may not be effective for large-scale problems. The second class of the methods was developed by Mangasarian and De Leone [18]. In their method, a decreasing sequence $\mu^k \to 0$ is stipulated, and for each μ^k an approximate solution $x(\mu^k)$ is computed by applying SOR algorithms to the dual problem of (1.3). They showed that if the residual inaccuracy of $x(\mu^k)$ falls below a certain threshold related to μ^k, the approximate sequence $\{x(\mu^k)\}$ converges to the least 2-norm solution as $k \to \infty$. It is worth mentioning that Kiwiel [7] extended the method in [18] to piecewise linear programs which include the linear program as a special case.

In summary, the aforementioned approaches focus on solving the problem (1.3) or its dual problem. We may categorize them as (sequential) quadratic programming methods for solving the least 2-norm solution of a linear program. Of course, in addition to SOR methods, the problem (1.3) or its dual problem may also be solved by other algorithms, see Lin and Pang [10] and the references therein.

Besides the methods using an equivalent perturbed quadratic program, the least 2-norm solution of a linear program can also be obtained by solving an equivalent unconstrained convex minimization problem. The first result in this aspect was due to Mangasarian [15]. In [15], Mangasarian first proved that the least 2-norm solution problem of linear programs can be transformed into an equivalent unconstrained minimization of a parameter-free convex continuously differentiable function. As a result, some unconstrained optimization methods can be used to solve the least 2-norm solution of a linear program. Recently, Kanzow et al [5] studied another equivalent unconstrained reformulation of the least 2-norm solution problem. Their method is based on the result of Smith and Wolkowicz [24] which is essentially related to the result of Mangasarian and Meyer (Corollary 2 in [19]). Based on their reformulation, Kanzow, Qi and Qi [5] proposed a Newton-type method to solve their unconstrained minimization problem. However, unlike Mangasarian's reformulation, the unconstrained minimization problem in Kanzow, Qi and Qi [5] contains a parameter which is required to be sufficiently large (but it is unknown in advance). Also, their convergence analysis needs certain relatively restrictive assumptions such as the strict feasibility of (1.1) and the nondegeneracy of the least 2-norm solution.

It is well-known that a linear program can also be formulated as an equivalent linear complementarity problem (LCP). In fact, writing out the Karush-Kuhn-Tucker (KKT) optimality conditions of the linear program (1.1), we have

$$\left\{\begin{array}{l} s + A^T y - c = 0, \\ z - Ax + b = 0, \\ (x, y, s, z) \geq 0, \;\; x^T s = y^T z = 0, \end{array}\right. \tag{1.4}$$

which can be written as the following monotone LCP:

$$\begin{bmatrix} s \\ z \end{bmatrix} = \begin{bmatrix} O & -A^T \\ A & O \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} c \\ -b \end{bmatrix} \geq 0, \begin{bmatrix} x \\ y \end{bmatrix} \geq 0, \begin{bmatrix} s \\ z \end{bmatrix}^T \begin{bmatrix} x \\ y \end{bmatrix} = 0. \tag{1.5}$$

Thus, locating the least 2-norm solutions of primal and dual linear programs is completely equivalent to finding a least 2-norm solution of the above monotone LCP. For LCPs, the least 2-norm solution has also been extensively studied by several authors, for example, Subramanian [26], Mangasarian [17], Sznajder and Gowda [27], and Zhao and Li [35, 36, 37]. The least 2-norm solution of a complementarity problem is also related to Tikhonov regularization methods for complementarity problems (see for instance, Isac [4], Facchinei [2], Facchinei and Kanzow [3]). In fact, in [26, 27, 35], it is shown that the Tikhonov regularization trajectory of a monotone complementarity problem converges to the least 2-norm solution of the problem. In [35], a new homotopy continuation trajectory, later called regularized central path in [37], is constructed for complementarity problems. It turns out that this new trajectory converges to the least 2-norm solution of a monotone complementarity problem as the parameter approaches to zero.

Motivated by recent results in [35, 36, 37], the purpose of this paper is to develop a new theory and an alternative computational method for the least 2-norm solution of linear programs. The proposed method is different from most of the existing methods which either require additional conditions besides the solvability of the problem, or have to solve quadratic programs successively. The proposed algorithm in this paper does not impose any assumption on the problem. It is convergent regardless of whether the linear program is solvable or not. If the problem (1.1) is solvable, then the iterates generated by the proposed algorithm converge to the least 2-norm solution. If the problem has no solution, the iterates still converge to a point which gives a minimal KKT residual of the problem (1.1). This algorithm is a kind of interior-point-like path-following algorithm (but not an interior-point algorithm), which is based on a new concept of regularized central path $\{x(\mu) : \mu > 0\}$ of a linear program. Remarkable features of this path are that its existence and convergence for any (solvable or unsolvable) linear program can be guaranteed. These features distinguish it from the conventional central path whose existence and boundedness require that the primal and the dual have interior-points, which in turn implies that both primal and dual problems have bounded solution sets (see, Theorem 5.10.1 and Corollary of Theorem 3.4.1 in [23]). When a linear program has an unbounded solution set in which case the problem is unstable (see, Robinson [22]), the interior-point does not exist, and hence the central path does not exist. However, the regularized central path proposed in this paper always exists for any linear program, and converges, as μ tends to zero, to the least 2-norm solution of any solvable linear program despite of the unboundedness of its solution set. This motivates us to design a new path-following method for linear programs. To our knowledge, the proposed method can be viewed as the first (interior-point-like) path-following algorithm for the least 2-norm solution of a linear program.

In the next section, we introduce the concept of a regularized central path for linear programs. In Section 3, we specify a path-following algorithm. In Section 4, we prove the global convergence of the algorithm. The unsolvable case is studied in Section 5. Numerical results are illustrated in Section 6. Conclusions are given in the last section.

Throughout the paper, we use the standard notation found in the interior-point algorithm literature. For example, all the vectors are column. For vectors u and $v \in R^n$, we also use (u, v) to denote the column vector $(u^T, v^T)^T$ if there is no confusion. The vector e denotes the vector of ones, and its dimension, unless otherwise stated, depends on the context. For a vector x, x_+ denotes the vector with components $(x_+)_i = \max\{x_i, 0\}, i = 1, ..., n$, and X denotes the corresponding diagonal matrix, i.e., $X = \mathrm{diag}(x)$. R^n_+ denotes the nonnegative orthant of n-dimensional Euclidean space R^n. If $x \in R^n_+$, we also write it as $x \geq 0$. In particular, $x > 0$ means that all components of x are positive.

2. Regularized central path. We begin with recalling the concept of a central path of a linear program. The linear program (1.1) can be rewritten as

$$\min\{c^T x : Ax - z = b,\ (x, z) \geq 0\}.$$

The central path is defined by a parameter $\mu > 0$, and for each $\mu > 0$, it is the solution to the following logarithmic barrier problem:

$$\begin{array}{ll} \min & c^T x - \mu \left(\sum_{i=1}^n \log x_i + \sum_{i=1}^m \log z_i\right) \\ \text{s.t.} & Ax - z = b \\ & (x > 0, z > 0). \end{array}$$

The Lagrangian of the above problem is

$$L_\mu(x, y, z) = c^T x + y^T (z - Ax + b) - \mu \left(\sum_{i=1}^n \log x_i + \sum_{i=1}^m \log z_i \right), \tag{2.1}$$

where $y \in R^m$ is the Lagrange multiplier vector corresponding to the constraint $Ax - z = b$. Thus, the central path is actually defined by the stationary point of the above Lagrange function, that is,

$$0 = \nabla L_\mu(x, y, z) = \left(\frac{\partial L_\mu}{\partial x}, \frac{\partial L_\mu}{\partial y}, \frac{\partial L_\mu}{\partial z}\right) = \begin{pmatrix} c - A^T y - \mu X^{-1} e \\ z - Ax + b \\ y - \mu Z^{-1} e \end{pmatrix},$$

which, by setting $s = \mu X^{-1} e > 0$, can be written as

$$\begin{aligned} Xs &= \mu e, \\ Yz &= \mu e, \\ s + A^T y - c &= 0, \\ z - Ax + b &= 0, \\ (x, y, s, z) &> 0. \end{aligned}$$

It is well known that for every $\mu > 0$, the above system has a unique solution denoted by $(x(\mu), y(\mu), s(\mu), z(\mu))$ if and only if the primal and the dual problems have interior points. If the primal and the dual have interior points, $x(\mu)$ converges (as $\mu \to 0$) to the analytic center of the primal optimal face, and $y(\mu)$ converges to the dual optimal face (Theorems 5.10.1 and 5.10.3 in [23], or Theorems 2.16 and 2.17 in [34]). Clearly, the analytic center is not necessarily the least 2-norm solution.

It is worth pointing out that the existence of the central path is not guaranteed for the case when the problem has an unbounded optimal solution set, i.e., when the linear program has no interior point (see for instance, Theorem 5.10.1 and the corollary of Theorem 3.4.1 in [23]). We now construct a new smooth path that is expected to converge to the least 2-norm solution even when the problem has an unbounded solution set. We first define a perturbed Lagrange function of (2.1). Notice that in (2.1), the Lagrange multiplier y is related to the decision variable of the dual problem. In fact, let $\nabla L_\mu(x(\mu), y(\mu), z(\mu)) = 0$. If $(x(\mu), y(\mu), z(\mu)) \to (x^*, y^*, z^*)$ as $\mu \to 0$, then (x^*, y^*, z^*) satisfies the KKT system (1.4). By the theory of linear programming, y^* is an optimal solution to the dual problem. Thus, in order to obtain the least 2-norm solution of the primal and the dual linear programs, we consider the following augmented Lagrange function:

$$\begin{aligned} \mathcal{L}_{(\mu,\theta)}(x, y, z) :&= c^T x + y^T (z - Ax + b) - \mu \left(\sum_{i=1}^n \log x_i + \sum_{i=1}^m \log z_i \right) \\ &+ \frac{1}{2} \theta (\|x\|_2^2 - \|y\|_2^2), \end{aligned} \tag{2.2}$$

where μ and θ are two positive parameters. The term $\theta(\|x\|_2^2 - \|y\|_2^2)$ attached to the Lagrangian (2.1) is used to force the stationary point of the augmented Lagrange function to approach the least 2-norm solution. It will be seen from our later discussion that the above augmented form is a judicious choice for locating the least 2-norm solution, and for covering the aforementioned case of an unbounded solution set. Although the parameters μ and θ can be independent, for simplicity, however, we consider here only the case of $\theta = \mu^p$, where $p \in (0, 1)$ is a fixed constant. Thus, the above function (2.2) can be written as the following one-parameter form:

$$\begin{aligned} \Phi_\mu(x, y, z) :&= c^T x + y^T (z - Ax + b) - \mu \left(\sum_{i=1}^n \log x_i + \sum_{i=1}^m \log z_i \right) \\ &+ \frac{1}{2} \mu^p (\|x\|_2^2 - \|y\|_2^2). \end{aligned} \tag{2.3}$$

We are now ready to define the concept of a regularized central path. Analogous to the central path which is the stationary point of Lagrange function (2.1), the so-called *regularized central path* can be defined by the stationary point of the augmented Lagrange function (2.3), that is, $\nabla\Phi_\mu(x, y, z) = 0$. Thus, we have the following definition.

DEFINITION 2.1. *The curve* $\{(x(\mu), y(\mu), s(\mu), z(\mu)) : \mu > 0\}$ *is said to be a regularized central path if for each* $\mu > 0$, $(x(\mu), y(\mu), s(\mu), z(\mu))$ *is the solution to the following system:*

$$\begin{cases} Xs = \mu e, \\ Yz = \mu e, \\ s + A^T y - c = \mu^p x, \\ z - Ax + b = \mu^p y, \\ (x, y, s, z) > 0. \end{cases} \tag{2.4}$$

The set $\{(x(\mu), z(\mu)) : \mu > 0\}$ can be called the primal regularized central path, and $\{(y(\mu), s(\mu)) : \mu > 0\}$ the dual regularized central path. The following result states that the existence of the regularized central path can be ensured in all situations. This path converges to the unique least 2-norm solution as long as the linear program in question is solvable. Thus, the regularized central path provides us a novel and powerful solution scheme for linear programming problems.

THEOREM 2.1. *For any linear program (1.1), the following holds:*

(i) For each $\mu > 0$, *the system (2.4) has a unique solution* $(x(\mu), z(\mu), y(\mu), s(\mu)) > 0$.

(ii) For any finite number $0 < \hat\mu < \infty$, *the set* $\{(x(\mu), z(\mu), y(\mu), s(\mu)) : \mu \in (0, \hat\mu]\}$ *is bounded if and only if the linear problem (1.1) is solvable.*

(iii) The linear problem (1.1) is solvable if and only if $(x(\mu), z(\mu), y(\mu), s(\mu))$ *converges, as* $\mu \to 0$, *to* (x^*, z^*, y^*, s^*) *where* x^* *and* y^* *are least 2-norm solutions of the primal and the dual problems, respectively.*

Proof. It is evident that the system (2.4) can be written as

$$\begin{bmatrix} s \\ z \end{bmatrix} = \begin{bmatrix} \mu^p I & -A^T \\ A & \mu^p I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} c \\ -b \end{bmatrix} > 0, \quad \begin{bmatrix} x \\ y \end{bmatrix} > 0, \ U \begin{bmatrix} s \\ z \end{bmatrix} = \mu e, \tag{2.5}$$

where $e \in R^{m+n}$, and $U = \begin{bmatrix} X & O \\ O & Y \end{bmatrix}$. Denote

$$M = \begin{bmatrix} O & -A^T \\ A & O \end{bmatrix}, \quad u = \begin{bmatrix} x \\ y \end{bmatrix}, \quad v = \begin{bmatrix} s \\ z \end{bmatrix}, \quad q = \begin{bmatrix} c \\ -b \end{bmatrix}.$$

Then the system (2.5) can be further written as

$$v = Mu + q + \mu^p u > 0, \ u > 0, \ Uv = \mu e.$$

Under the one-to-one transformation of $\mu = \varepsilon/(1-\varepsilon)$ where $\varepsilon \in (0, 1)$, the above system is equivalent to

$$(1-\varepsilon)v = (1-\varepsilon)(Mu + q + \phi(\varepsilon)u) > 0, \ u > 0, \ (1-\varepsilon)Uv = \varepsilon e,$$

where $\phi(\varepsilon) = \left(\frac{\varepsilon}{1-\varepsilon}\right)^p$. Denote by $w = (1-\varepsilon)v$, The above system can be finally written as

$$\bar{\mathcal{H}}(u, w, \varepsilon) := \begin{pmatrix} Uw - \varepsilon e \\ w - (1-\varepsilon)(Mu + q + \phi(\varepsilon)u) \end{pmatrix} = 0, \ (u, w) > 0, \tag{2.6}$$

Noting that $p \in (0,1)$, we have $\varepsilon/\phi(\varepsilon) \to 0$ as $\varepsilon \to 0$. Since the matrix M is a monotone matrix, it must be a P_0 matrix or a P_* matrix. By Theorem 4.2(a) or Theorem 5.2(a) in [35] (but applied to monotone LCP), we conclude that the above system (2.6) has a unique solution $(u(\varepsilon), w(\varepsilon))$ for each given $\varepsilon > 0$. The result (i) is proved.

To see that (ii) holds, we first note that if the path $\{(x(\mu), y(\mu), s(\mu), z(\mu)) : \mu \in (0, \hat{\mu}]\}$ is bounded, taking $\mu \to 0$ in system (2.4) we see that any accumulation point of the path is a solution to KKT system (1.4), and thus it is a solution to the linear program. Conversely, assume that the linear program is solvable. This is equivalent to saying that the LCP (1.5) is solvable. Notice that (2.4) can be written as (2.5). It follows from Theorem 5.1(b) in [35] that the path $\{(x(\mu), y(\mu), s(\mu), z(\mu)) : \mu \in (0, \hat{\mu}]\}$ is bounded. Result (ii) holds. Since the solutions of LCP (1.5) are the same as the solutions of the primal and the dual programs (1.1) and (1.2), result (iii) is an immediate consequence of Theorem 5.2 in [35]. □

From the above result, we obtain the following characterization of the least 2-norm solution of a linear program.

COROLLARY 2.1. (x^*, y^*) *is the least 2-norm solution pair to the primal and the dual problems if and only if it is the unique limiting point of the regularized central path as* $\mu \to 0$. *Equivalently, if the problem (1.1) has no optimal solution, i.e., the problem (1.1) is unsolvable, if and only if the regularized central path is divergent to infinity as* $\mu \to 0$.

3. Algorithm. Our algorithm can tackle both solvable and unsolvable linear programming problems. For simplicity, however, we consider first the solvable problems. The general case, including unsolvable problems, is treated in Section 5.

For a fixed scalar $p \in (0,1)$, we denote $\mathcal{F}_\mu : R^{2(n+m)} \to R^{2(n+m)}$ by

$$\mathcal{F}_\mu(x, y, s, z) = \begin{pmatrix} Xs - \mu e \\ Yz - \mu e \\ s + A^T y - c - \mu^p x \\ z - Ax + b - \mu^p y \end{pmatrix}. \tag{3.1}$$

Note that the regularized central path is given by the following system:

$$\mathcal{F}_\mu(x, y, s, z) = 0, \ (x, y, s, z) > 0.$$

We also note that the vector (x^*, y^*, s^*, z^*) is a solution to the KKT system (1.4) if and only if it satisfies

$$\mathcal{F}_0(x^*, y^*, s^*, z^*) = 0, \ (x^*, y^*, s^*, z^*) \geq 0.$$

To give a path-following algorithm, we employ the following set as a neighborhood of the regularized central path:

$$\mathcal{N}_\beta(\mu) := \{(x, y, s, z) > 0 : \ \|\mathcal{F}_\mu(x, y, s, z)\|_\infty \leq \beta\mu\},$$

where $\beta \in (0,1)$ is a fixed scalar. From a starting point $(x^0, y^0, s^0, z^0) > 0$, the purpose of our path-following algorithm is to generate a positive sequence (x^k, y^k, s^k, z^k) confined in the above neighborhood. This sequence converges to a solution of the problem. In each step of the algorithm, only one linear algebraic equation is solved and the Armijo-type line search is used to determine the stepsize. While the iteration of this algorithm proceeds in the positive orthant, i.e., all the iterates maintain positivity, the iterates are not necessarily to be interior points of the problem. In fact, this algorithm does not require that the problem possess an interior point, and thus it does not belong to the class of central path-based interior-point algorithms.

ALGORITHM 3.1:

Step 1. (Initial step) Let $\beta \in (0,1)$ be a positive scalar. Assign scalars α_1, α_2, and σ in (0,1). Select $(x^0, y^0, s^0, z^0) > 0$ and $\mu^0 \in (0, \infty)$ such that $(x^0, y^0, s^0, z^0) \in \mathcal{N}_\beta(\mu^0)$.

Step 2. (Centering step) If $\mathcal{F}_{\mu^k}(x^k, y^k, s^k, z^k) = 0$, set

$$(x^{k+1}, y^{k+1}, s^{k+1}, z^{k+1}) = (x^k, y^k, s^k, z^k),$$

and go to Step 3. Otherwise, let $(\Delta x^k, \Delta y^k)$ be the solution to the following equation

$$\begin{bmatrix} S^k + (\mu^k)^p X^k & -X^k A^T \\ Y^k A & Z^k + (\mu^k)^p Y^k \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} \mu^k e - X^k s^k \\ \mu^k e - Y^k z^k \end{bmatrix} - \begin{bmatrix} X^k(-s^k + (\mu^k)^p x^k - A^T y^k + c) \\ Y^k(-z^k + Ax^k + (\mu^k)^p y^k - b) \end{bmatrix}. \tag{3.2}$$

Then, set

$$\begin{bmatrix} \Delta s^k \\ \Delta z^k \end{bmatrix} = \begin{bmatrix} (\mu^k)^p I & -A^T \\ A & (\mu^k)^p I \end{bmatrix} \begin{bmatrix} \Delta x^k \\ \Delta y^k \end{bmatrix} + \begin{bmatrix} -s^k + (\mu^k)^p x^k - A^T y^k + c \\ -z^k + Ax^k + (\mu^k)^p y^k - b \end{bmatrix}. \tag{3.3}$$

Let

$$\begin{aligned} \bar{\alpha} = \arg\max\{\alpha > 0 : \; & x^k + \lambda \Delta x^k > 0,\ y^k + \lambda \Delta y^k > 0,\ s^k + \lambda \Delta z^k > 0, \\ & z^k + \lambda \Delta z^k > 0 \text{ for all } \lambda \in (0, \alpha]\}. \end{aligned}$$

Let λ_k be the maximum among the values of $\bar{\alpha}, \alpha_1 \bar{\alpha}, \alpha_1^2 \bar{\alpha}, ...$ such that

$$\begin{aligned} & \|\mathcal{F}_{\mu^k}(x^k + \lambda_k \Delta x^k, y^k + \lambda_k \Delta y^k, s^k + \lambda_k \Delta s^k, z^k + \lambda_k \Delta z^k)\|_\infty \\ & \le (1 - \sigma \lambda_k) \|\mathcal{F}_{\mu^k}(x^k, y^k, s^k, z^k)\|_\infty. \end{aligned} \tag{3.4}$$

Set

$$(x^{k+1}, y^{k+1}, s^{k+1}, z^{k+1}) = (x^k, y^k, s^k, z^k) + \lambda_k (\Delta x^k, \Delta y^k, \Delta s^k, \Delta z^k),$$

and go to Step 3.

Step 3. (Reduction step for μ) Let γ^k be the maximum among the values of $\alpha_2, \alpha_2^2, ...$ such that

$$(x^{k+1}, y^{k+1}, s^{k+1}, z^{k+1}) \in \mathcal{N}_\beta((1 - \gamma^k)\mu^k),$$

i.e.,

$$\|\mathcal{F}_{(1-\gamma^k)\mu^k}(x^{k+1}, y^{k+1}, s^{k+1}, z^{k+1})\|_\infty \le \beta(1 - \gamma^k)\mu^k.$$

Set $\mu^{k+1} := (1 - \gamma^k)\mu^k$, and go to Step 2.

Remark 3.1. In numerical implementation, the initial points and some stopping criterion are needed. For the above algorithm, we may use $\|\mathcal{F}_0(x^k, y^k, s^k, z^k)\|_\infty < \varepsilon$ or $\mu^k < \varepsilon$ as the stopping criterion, where ε is a termination tolerance. The initial point for the above algorithm can be constructed without any additional cost. For instance, a practical initial step proceeds as follows:

Initial Step: Let $(x^0, y^0) = e$ $(\in R^{n+m})$. *Choose* μ^0 *such that*

$$\mu^0 > \max\left\{1, \left\| \begin{pmatrix} A^T y^0 - c \\ -Ax^0 + b \end{pmatrix} \right\|_\infty \right\}.$$

Let $(s^0, z^0) = (\mu^0)^p e$ $(\in R^{n+m})$, *and let*

$$\eta := \frac{\|\mathcal{F}_{\mu^0}(x^0, y^0, s^0, z^0)\|_\infty}{\mu^0}.$$

Then, assign $\beta \in [\eta, 1)$.

From the above choice, by (3.1) we see that

$$\|\mathcal{F}_{\mu^0}(x^0, y^0, s^0, z^0)\|_\infty = \max\left\{ \left|\mu^0 - (\mu^0)^p\right|, \left\| \begin{pmatrix} A^T y^0 - c \\ -Ax^0 + b \end{pmatrix} \right\|_\infty \right\}.$$

By the choice of μ^0, it follows that $0 < \eta < 1$. Thus, $\eta \le \beta < 1$ and $(x^0, y^0, s^0, z^0) \in \mathcal{N}_\beta(\mu^0)$.

Remark 3.2. We now point out that at the current point $(x^k, y^k, s^k, z^k) > 0$ the vector $(\Delta x^k, \Delta y^k, \Delta s^k, \Delta z^k)$ determined by systems (3.2) and (3.3) is unique. In fact, it is easy to see that $(\Delta x^k, \Delta y^k, \Delta s^k, \Delta z^k)$ is a solution to the systems (3.2) and (3.3) if and only if it is the solution to the following system:

$$\left\{ \begin{array}{rcl} S^k \Delta x + X^k \Delta s &=& \mu^k e - X^k s^k, \\ Z^k \Delta y + Y^k \Delta z &=& \mu^k e - Y^k z^k, \\ \Delta s - (\mu^k)^p \Delta x + A^T \Delta y &=& -s^k - A^T y^k + (\mu^k)^p x^k + c, \\ \Delta z - A \Delta x + (\mu^k)^p \Delta y &=& -z^k + Ax^k + (\mu^k)^p y^k - b. \end{array} \right. \tag{3.5}$$

which is a $2(m+n)$-dimensional linear system. Notice that the Jacobian Matrix of $\mathcal{F}_{\mu^k}(x, y, s, z)$ at $(x^k, y^k, s^k, z^k) > 0$ is given by

$$\nabla \mathcal{F}_{\mu^k}(x^k, y^k, s^k, z^k) = \begin{bmatrix} S^k & O & X^k & O \\ O & Z^k & O & Y^k \\ -(\mu^k)^p I & A^T & I & O \\ -A & -(\mu^k)^p I & O & I \end{bmatrix}. \tag{3.6}$$

The system (3.5) coincides with the following:

$$\mathcal{F}_{\mu^k}(x^k, y^k, s^k, z^k) + \nabla \mathcal{F}_{\mu^k}(x^k, y^k, s^k, z^k)(\Delta x, \Delta y, \Delta s, \Delta z) = 0. \tag{3.7}$$

Hence, the direction $(\Delta x, \Delta y, \Delta s, \Delta z)$ is actually the Newton direction determined by (3.7). Since the matrix

$$\begin{bmatrix} (\mu^k)^p I & -A^T \\ A & (\mu^k)^p I \end{bmatrix}$$

is positive semidefinite for any $\mu^k > 0$, at the positive point (x^k, y^k, s^k, z^k) the Jacobian matrix $\nabla \mathcal{F}_{\mu^k}(x^k, y^k, s^k, z^k)$ given by (3.6) is nonsingular. This fact follows from Lemma 5.4 in Kojima et al [9]. Thus the system (3.7) has a unique solution, and hence the systems (3.2) and (3.3) have a unique solution. This can also be explained another way. In fact, at $(x^k, y^k, s^k, z^k) > 0$, it is easy to verify that the nonsingularity of the matrix $\nabla \mathcal{F}_{\mu^k}(x^k, y^k, s^k, z^k)$ implies the nonsingularity of the matrix

$$\begin{bmatrix} S^k + (\mu^k)^p X^k & -X^k A^T \\ Y^k A & Z^k + (\mu^k)^p Y^k \end{bmatrix}.$$

While the system (3.2) together with (3.3) is equivalent to the system (3.5) or (3.7), we choose to solve the system (3.2) since it has lower dimension than (3.5).

4. Global convergence. In this section, we show that whenever the solution set is nonempty, the iterates $\{(x^k, y^k)\}$ generated by Algorithm 3.1 converge to the least 2-norm solutions of the primal and the dual linear programs. We first show that the algorithm is well-defined.

LEMMA 4.1. *Algorithm 3.1 is well-defined. The sequence $\{\mu^k\}$ is monotonically decreasing, and $(x^k, y^k, s^k, z^k) \in \mathcal{N}_\beta(\mu^k)$ for all $k \geq 0$.*

Proof. We verify that each step of the algorithm is well-defined. By Remark 3.1, the first step is well-defined. The starting point satisfies

$$(x^0, y^0, s^0, z^0) > 0, \ (x^0, y^0, s^0, z^0) \in \mathcal{N}_\beta(\mu^0).$$

By induction, we now assume that

$$(x^k, y^k, s^k, z^k) > 0, \ (x^k, y^k, s^k, z^k) \in \mathcal{N}_\beta(\mu^k).$$

We show that the next iterate $(x^{k+1}, y^{k+1}, s^{k+1}, z^{k+1})$ generated by the algorithm still maintains positivity, and satisfies the condition $(x^{k+1}, y^{k+1}, s^{k+1}, z^{k+1}) \in \mathcal{N}_\beta(\mu^{k+1})$. By the positivity of (x^k, y^k, s^k, z^k), from Remark 3.2, the system defined by (3.2) and (3.3) has a unique solution, and the Newton direction $(\Delta x^k, \Delta y^k, \Delta s^k, \Delta z^k)$ is a descent direction of the function $\|\mathcal{F}_{\mu^k}(x, y, s, z)\|_\infty$ at the current point $(x^k, y^k, s^k, z^k) > 0$. Thus, the line search rule (3.4) is well-defined, and hence Step 2 is well-defined. Since

$$\|\mathcal{F}_{\mu^k}(x^k, y^k, s^k, z^k)\|_\infty \leq \beta\mu^k \text{ and } 1 - \sigma\lambda_k < 1,$$

from (3.4) we have

$$\|\mathcal{F}_{\mu^k}(x^k + \lambda_k\Delta x^k, y^k + \lambda_k\Delta y^k, s^k + \lambda_k\Delta s^k, z^k + \lambda_k\Delta z^k)\|_\infty \leq \beta\mu^k,$$

that is,

$$\|\mathcal{F}_{\mu^k}(x^{k+1}, y^{k+1}, s^{k+1}, z^{k+1})\|_\infty \leq \beta\mu^k,$$

which implies that

$$\|X^{k+1}s^{k+1} - \mu^k e\|_\infty \leq \beta\mu^k, \ \|Y^{k+1}z^{k+1} - \mu^k e\|_\infty \leq \beta\mu^k.$$

By the choice of $\bar{\alpha}$ and λ_k, we see that $(x^{k+1}, y^{k+1}, s^{k+1}, z^{k+1})$ is nonnegative. Combining this fact and the above inequalities, where $0 < \beta < 1$, concludes that the next iterate $(x^{k+1}, y^{k+1}, s^{k+1}, z^{k+1})$ must be positive.

We now show that Step 3 is well-defined, and hence the next iterate is contained in the set $\mathcal{N}_\beta(\mu^{k+1})$. There are two possible cases.

Case 1: $\mathcal{F}_{\mu^k}(x^k, y^k, s^k, z^k) = 0$. According to the construction of the algorithm, $(x^{k+1}, y^{k+1}, s^{k+1}, z^{k+1}) = (x^k, y^k, s^k, z^k)$. By continuity, there is a γ^k determined by Step 3 such that

$$\|\mathcal{F}_{(1-\gamma^k)\mu^k}(x^{k+1}, y^{k+1}, s^{k+1}, z^{k+1})\|_\infty \leq \beta(1 - \gamma^k)\mu^k.$$

Thus, by setting $\mu^{k+1} = (1 - \gamma^k)\mu^k$, we have $(x^{k+1}, y^{k+1}, s^{k+1}, z^{k+1}) \in \mathcal{N}_\beta(\mu^{k+1})$.

Case 2: $\mathcal{F}_{\mu^k}(x^k, y^k, s^k, z^k) \neq 0$. In the case, the next point $(x^{k+1}, y^{k+1}, s^{k+1}, z^{k+1})$ is determined by (3.4). We now show that $(x^{k+1}, y^{k+1}, s^{k+1}, z^{k+1}) \in \mathcal{N}_\beta(\mu^{k+1})$ still holds. For any $(x, y, s, z) > 0$ and $t_2 \geq t_1 \geq 0$, it is easy to verify that

$$\|\mathcal{F}_{t_1}(x, y, s, z) - \mathcal{F}_{t_2}(x, y, s, z)\|_\infty \leq t_2 - t_1 + (t_2^p - t_1^p)\|(x, y)\|_\infty.$$

Thus, by (3.4) and the above inequality, we have

$$
\begin{aligned}
&\|\mathcal{F}_{(1-\gamma)\mu^k}(x^{k+1}, y^{k+1}, s^{k+1}, z^{k+1})\|_\infty \\
&\le \|\mathcal{F}_{(1-\gamma)\mu^k}(x^{k+1}, y^{k+1}, s^{k+1}, z^{k+1}) - \mathcal{F}_{\mu^k}(x^{k+1}, y^{k+1}, s^{k+1}, z^{k+1})\|_\infty \\
&\quad + \|\mathcal{F}_{\mu^k}(x^{k+1}, y^{k+1}, s^{k+1}, z^{k+1})\|_\infty \\
&\le \gamma\mu^k + (\mu^k)^p[1-(1-\gamma)^p]\|(x^{k+1}, y^{k+1})\|_\infty + (1-\sigma\lambda_k)\|\mathcal{F}_{\mu^k}(x^k, y^k, s^k, z^k)\|_\infty \\
&\le \gamma\mu^k + (\mu^k)^p[1-(1-\gamma)^p]\|(x^{k+1}, y^{k+1})\|_\infty + (1-\sigma\lambda_k)\beta\mu^k \\
&= \left[\frac{\gamma + (\mu^k)^{p-1}[1-(1-\gamma)^p]\|(x^{k+1}, y^{k+1})\|_\infty}{(1-\gamma)\beta} + \frac{(1-\sigma\lambda_k)}{1-\gamma}\right]\beta(1-\gamma)\mu^k.
\end{aligned}
$$

Since $1-\sigma\lambda_k < 1$, there is a positive scalar $\hat{\gamma} > 0$ such that for all $\gamma \in (0, \hat{\gamma}]$, the term in the above bracket is less than one. Thus, for all sufficiently small $\gamma > 0$ we have

$$\|\mathcal{F}_{(1-\gamma)\mu^k}(x^{k+1}, y^{k+1}, s^{k+1}, z^{k+1})\|_\infty \le \beta(1-\gamma)\mu^k.$$

Step 3 is well-defined. Of course, the sequence $\{\mu^k\}$ is monotonically decreasing since $\mu^{k+1} = (1-\gamma^k)\mu^k$. ☐

By Lemma 4.1, the sequence $(x^k, y^k, s^k, z^k) \in \mathcal{N}_\beta(\mu^k)$ for all k, i.e.,

$$\|\mathcal{F}_{\mu^k}(x^k, y^k, s^k, z^k)\|_\infty \le \beta\mu^k \text{ and } (x^k, y^k, s^k, z^k) > 0. \tag{4.1}$$

We employ auxiliary sequences $(u^k, v^k, w^k, q^k) \in R^{2(n+m)}$ defined by

$$(u^k, v^k, w^k, q^k) = \frac{1}{\mu^k}\mathcal{F}_{\mu^k}(x^k, y^k, s^k, z^k). \tag{4.2}$$

Clearly, the above sequence $\{(u^k, v^k, w^k, q^k)\}$ is uniformly bounded. Indeed, combining (4.1) and (4.2) yields $\|(u^k, v^k, w^k, q^k)\|_\infty \le \beta$. The equation (4.2) can be written as

$$X^k s^k = \mu^k(e + u^k), \tag{4.3}$$

$$Y^k z^k = \mu^k(e + v^k), \tag{4.4}$$

$$s^k = -A^T y^k + c + (\mu^k)^p x^k + \mu^k w^k, \tag{4.5}$$

$$z^k = Ax^k - b + (\mu^k)^p y^k + \mu^k q^k. \tag{4.6}$$

These relations play a key role in the remainder analysis. We now prove the main result of this section.

THEOREM 4.1. *Assume that the solution set of linear program (1.1) is nonempty. The sequence (x^k, y^k, s^k, z^k) generated by Algorithm 3.1 converges to $(\hat{x}, \hat{y}, \hat{s}, \hat{z})$ where $\hat{x}$ is the least 2-norm solution of the primal linear program (1.1), and $\hat{y}$ is the least 2-norm solution of the dual problem (1.2).*

Proof. We prove this result in three steps:

(i) If the solution set is nonempty, then the iterative sequence (x^k, y^k, s^k, z^k) generated by the algorithm is bounded.

(ii) $\mu^k \to 0$ and $\|\mathcal{F}_{\mu^k}(x^k, y^k, s^k, z^k)\|_\infty \to 0$. Thus, every accumulation point of the iterative sequence is a solution to the linear program.

(iii) The accumulation point is unique and must be the least 2-norm solution of the linear program.

We now prove (i). Let x^* be an arbitrary optimal solution of (1.1), and y^* be an arbitrary optimal solution of its dual problem (1.2). Let (s^*, z^*) be given by

$$\begin{bmatrix} s^* \\ z^* \end{bmatrix} = \begin{bmatrix} O & -A^T \\ A & O \end{bmatrix} \begin{bmatrix} x^* \\ y^* \end{bmatrix} + \begin{bmatrix} c \\ -b \end{bmatrix}. \tag{4.7}$$

Then, it is easy to see that $(x^*, y^*, s^*, z^*) \geq 0$ and $(x^*)^T s^* = 0$ and $(y^*)^T z^* = 0$. That is, (x^*, y^*, s^*, z^*) satisfies the KKT system (1.4). It follows from (4.3) and (4.4) that

$$(x^k)^T s^k = \mu^k (n + e^T u^k), \ (y^k)^T z^k = \mu^k (m + e^T v^k). \tag{4.8}$$

Notice that for any $(x, y) \in R^{n+m}$, we have

$$\begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} (\mu^k)^p I & -A^T \\ A & (\mu^k)^p I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = (\mu^k)^p \left\| \begin{bmatrix} x \\ y \end{bmatrix} \right\|_2^2. \tag{4.9}$$

By the positivity of (x^k, y^k, s^k, z^k), (4.5), (4.6), (4.9), (4.8), and (4.7), we have

$$\begin{aligned}
0 &\leq \begin{bmatrix} x^* \\ y^* \end{bmatrix}^T \begin{bmatrix} s^k \\ z^k \end{bmatrix} + \begin{bmatrix} s^* \\ z^* \end{bmatrix}^T \begin{bmatrix} x^k \\ y^k \end{bmatrix} \\
&= \begin{bmatrix} x^* - x^k \\ y^* - y^k \end{bmatrix}^T \begin{bmatrix} s^k \\ z^k \end{bmatrix} + \begin{bmatrix} s^* \\ z^* \end{bmatrix}^T \begin{bmatrix} x^k \\ y^k \end{bmatrix} + \begin{bmatrix} x^k \\ y^k \end{bmatrix}^T \begin{bmatrix} s^k \\ z^k \end{bmatrix} \\
&= \begin{bmatrix} x^* - x^k \\ y^* - y^k \end{bmatrix}^T \left(\begin{bmatrix} (\mu^k)^p I & -A^T \\ A & (\mu^k)^p I \end{bmatrix} \begin{bmatrix} x^k \\ y^k \end{bmatrix} + \begin{bmatrix} c \\ -b \end{bmatrix} + \mu^k \begin{bmatrix} w^k \\ q^k \end{bmatrix} \right) \\
&\quad + \begin{bmatrix} s^* \\ z^* \end{bmatrix}^T \begin{bmatrix} x^k \\ y^k \end{bmatrix} + \begin{bmatrix} x^k \\ y^k \end{bmatrix}^T \begin{bmatrix} s^k \\ z^k \end{bmatrix} \\
&= - \begin{bmatrix} x^k - x^* \\ y^k - x^* \end{bmatrix}^T \begin{bmatrix} (\mu^k)^p I & -A^T \\ A & (\mu^k)^p I \end{bmatrix} \begin{bmatrix} x^k - x^* \\ y^k - x^* \end{bmatrix} \\
&\quad - \begin{bmatrix} x^k - x^* \\ y^k - y^* \end{bmatrix}^T \left(\begin{bmatrix} (\mu^k)^p I & -A^T \\ A & (\mu^k)^p I \end{bmatrix} \begin{bmatrix} x^* \\ y^* \end{bmatrix} + \begin{bmatrix} c \\ -b \end{bmatrix} + \mu^k \begin{bmatrix} w^k \\ q^k \end{bmatrix} \right) \\
&\quad + \begin{bmatrix} s^* \\ z^* \end{bmatrix}^T \begin{bmatrix} x^k \\ y^k \end{bmatrix} + \begin{bmatrix} x^k \\ y^k \end{bmatrix}^T \begin{bmatrix} s^k \\ z^k \end{bmatrix} \\
&= -(\mu^k)^p \left\| \begin{bmatrix} x^k - x^* \\ y^k - y^* \end{bmatrix} \right\|_2^2 + \begin{bmatrix} s^* \\ z^* \end{bmatrix}^T \begin{bmatrix} x^k \\ y^k \end{bmatrix} + \begin{bmatrix} x^k \\ y^k \end{bmatrix}^T \begin{bmatrix} s^k \\ z^k \end{bmatrix} \\
&\quad - \begin{bmatrix} x^k - x^* \\ y^k - y^* \end{bmatrix}^T \left((\mu^k)^p \begin{bmatrix} x^* \\ y^* \end{bmatrix} + \begin{bmatrix} s^* \\ z^* \end{bmatrix} + \mu^k \begin{bmatrix} w^k \\ q^k \end{bmatrix} \right) \\
&= -(\mu^k)^p \left\| \begin{bmatrix} x^k - x^* \\ y^k - y^* \end{bmatrix} \right\|_2^2 - (\mu^k)^p \begin{bmatrix} x^k - x^* \\ y^k - y^* \end{bmatrix}^T \begin{bmatrix} x^* + (\mu^k)^{1-p} w^k \\ y^* + (\mu^k)^{1-p} q^k \end{bmatrix} \\
&\quad + \begin{bmatrix} x^k \\ y^k \end{bmatrix}^T \begin{bmatrix} s^k \\ z^k \end{bmatrix} \\
&= -(\mu^k)^p \left\| \begin{bmatrix} x^k - x^* \\ y^k - y^* \end{bmatrix} \right\|_2^2 - (\mu^k)^p \begin{bmatrix} x^k - x^* \\ y^k - y^* \end{bmatrix}^T \begin{bmatrix} x^* + (\mu^k)^{1-p} w^k \\ y^* + (\mu^k)^{1-p} q^k \end{bmatrix} \\
&\quad + \mu^k (m + n + e^T u^k + e^T v^k).
\end{aligned} \tag{4.10}$$

Dividing both sides of the above by $(\mu^k)^p$, we have

$$\left\| \begin{bmatrix} x^k - x^* \\ y^k - y^* \end{bmatrix} \right\|_2^2 \le \left\| \begin{bmatrix} x^k - x^* \\ y^k - y^* \end{bmatrix} \right\|_2 \left\| \begin{bmatrix} x^* + (\mu^k)^{1-p} w^k \\ y^* + (\mu^k)^{1-p} q^k \end{bmatrix} \right\|_2 + \left(\mu^k\right)^{1-p}\left(m + n + e^T u^k + e^T v^k\right). \tag{4.11}$$

Since $\mu^k \le \mu^0$ and (u^k, v^k, w^k, q^k) is uniformly bounded, the boundedness of the iterative sequence (x^k, y^k, s^k, z^k) follows from the above inequality. The part (i) is proven.

We now prove part (ii). Since all iterates are confined in $\mathcal{N}_\beta(\mu^k)$, it implies that (4.1) holds for all k. Thus, to show $\|\mathcal{F}_{\mu^k}(x^k, y^k, s^k, z^k)\|_\infty \to 0$, it suffices to show that $\mu^k \to 0$. In fact, μ^k is monotonically decreasing since $\mu^{k+1} = (1 - \gamma^k)\mu^k$. Thus, there exists a scalar $\hat{\mu} \ge 0$ such that $\mu^k \to \hat{\mu}$. By (i), the sequence (x^k, y^k, s^k, z^k) is bounded. Without loss of generality, we may assume that $(x^k, y^k, s^k, z^k) \to (\hat{x}, \hat{y}, \hat{s}, \hat{z})$. Taking the limit in (4.1), we have that

$$\|\mathcal{F}_{\hat{\mu}}(\hat{x}, \hat{y}, \hat{s}, \hat{z})\|_\infty \le \beta\hat{\mu}, \ (\hat{x}, \hat{y}, \hat{s}, \hat{z}) \ge 0. \tag{4.12}$$

We assume to the contrary that $\hat{\mu} \ne 0$, i.e., $\hat{\mu} > 0$. We now derive a contradiction. Combining the fact $\mu^{k+1} = (1 - \gamma^k)\mu^k$ and $\mu^k \to \hat{\mu} > 0$ implies that $\gamma^k \to 0$ as $k \to \infty$.

We deduce from (4.12) that

$$\|\hat{X}\hat{s} - \hat{\mu}e\|_\infty \le \beta\hat{\mu}, \ \|\hat{Y}\hat{z} - \hat{\mu}e\|_\infty \le \beta\hat{\mu}.$$

Since $0 < \beta < 1$ and $(\hat{x}, \hat{y}, \hat{s}, \hat{z}) \ge 0$, the above inequality implies that $(\hat{x}, \hat{y}, \hat{s}, \hat{z}) > 0$. Thus, by Remark 3.2, the Jacobian $\nabla\mathcal{F}_{\hat{\mu}}(\hat{x}, \hat{y}, \hat{s}, \hat{z})$ is nonsingular, and hence the matrix

$$\begin{bmatrix} \hat{S} + \hat{\mu}^p\hat{X} & -\hat{X}A^T \\ \hat{Y}A & \hat{Z} + \hat{\mu}^p\hat{Y} \end{bmatrix}$$

is nonsingular. Therefore, the following system has a unique solution, denoted by $(\Delta\hat{x},\, \Delta\hat{y},\, \Delta\hat{s},\, \Delta\hat{z})$,

$$\begin{bmatrix} \hat{S} + \hat{\mu}^p\hat{X} & -\hat{X}A^T \\ \hat{Y}A & \hat{Z} + \hat{\mu}^p\hat{Y} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} \hat{\mu}e - \hat{X}\hat{s} \\ \hat{\mu}e - \hat{Y}\hat{z} \end{bmatrix} - \begin{bmatrix} \hat{X}(-\hat{s} + \hat{\mu}^p\hat{x} - A^T\hat{y} + c) \\ \hat{Y}(-\hat{z} + A\hat{x} + \hat{\mu}^p\hat{y} - b) \end{bmatrix},$$

$$\begin{bmatrix} \Delta s \\ \Delta z \end{bmatrix} = \begin{bmatrix} \hat{\mu}^p I & -A^T \\ A & \hat{\mu}^p I \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} + \begin{bmatrix} -\hat{s} + \hat{\mu}^p\hat{x} - A^T\hat{y} + c \\ -\hat{z} + A\hat{x} + \hat{\mu}^p\hat{y} - b \end{bmatrix}.$$

By Remark 3.2, this is equivalent to

$$\mathcal{F}_{\hat{\mu}}(\hat{x}, \hat{y}, \hat{s}, \hat{z}) + \nabla\mathcal{F}_{\hat{\mu}}(\hat{x}, \hat{y}, \hat{s}, \hat{z})(\Delta\hat{x}, \Delta\hat{y}, \Delta\hat{s}, \Delta\hat{z}) = 0,$$

which implies that $(\Delta\hat{x}, \Delta\hat{y}, \Delta\hat{s}, \Delta\hat{z})$ is a Newton descent direction of $\|\mathcal{F}_{\hat{\mu}}(x, y, s, z)\|_\infty$ at $(\hat{x}, \hat{y}, \hat{s}, \hat{z})$. Thus the linear search stepsize $\hat{\lambda}$ in (3.4) and $\hat{\gamma}$ in Step 3 of Algorithm 3.1 are both bounded below by a positive constant. By continuity, it follows that

$$(\Delta x^k, \Delta y^k, \Delta s^k, \Delta z^k, \mu^k, \lambda^k, \gamma^k) \to (\Delta\hat{x}, \Delta\hat{y}, \Delta\hat{s}, \Delta\hat{z}, \hat{\mu}, \hat{\lambda}, \hat{\gamma}).$$

In particular, $\gamma^k \to \hat{\gamma} > 0$, which contradicts $\gamma^k \to 0$. This contradiction shows that μ^k must converge to zero and thus $\|\mathcal{F}_{\mu^k}(x^k, y^k, s^k, z^k)\| \to 0$ as $k \to \infty$. Therefore, for any accumulation point $(\hat{x}, \hat{y}, \hat{s}, \hat{z})$, by continuity we have $\|\mathcal{F}_0(\hat{x}, \hat{y}, \hat{s}, \hat{z})\| = 0$ which implies that $(\hat{x}, \hat{y})$ is a solution pair to the primal and the dual linear programs.

Finally, we show that the accumulation point of the iterates is the unique least 2-norm solution. From (4.10), we have

$$\left\| \begin{bmatrix} x^k - x^* \\ y^k - y^* \end{bmatrix} \right\|_2^2 \le - \begin{bmatrix} x^k - x^* \\ y^k - y^* \end{bmatrix}^T \begin{bmatrix} x^* + (\mu^k)^{1-p} w^k \\ y^* + (\mu^k)^{1-p} q^k \end{bmatrix} + (\mu^k)^{1-p}(m + n + e^T u^k + e^T v^k).$$

Let $(\hat{x}, \hat{y}, \hat{s}, \hat{z})$ be an arbitrary accumulation point of the iterates. Notice that $p \in (0, 1)$, $\mu^k \to 0$, and (u^k, v^k, w^k, q^k) is bounded. Taking the limit in the above inequality, we have

$$\left\| \begin{bmatrix} \hat{x} - x^* \\ \hat{y} - y^* \end{bmatrix} \right\|_2^2 \le - \begin{bmatrix} \hat{x} - x^* \\ \hat{y} - y^* \end{bmatrix}^T \begin{bmatrix} x^* \\ y^* \end{bmatrix},$$

which can be written as

$$\left\| \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} \right\|_2^2 \le \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix}^T \begin{bmatrix} x^* \\ y^* \end{bmatrix} \le \left\| \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} \right\|_2 \left\| \begin{bmatrix} x^* \\ y^* \end{bmatrix} \right\|_2.$$

Since (x^*, y^*) is an arbitrary solution pair of the primal and the dual, from the above inequality we deduce that $\hat{x}$ and $\hat{y}$ are the least 2-norm solutions of the primal and the dual, respectively. (In fact, substituting (x^*, y^*) by $(x^*, \hat{y})$ and $(\hat{x}, y^*)$, respectively, we see that the above inequality implies that $\|(\hat{x}, \hat{y})\|_2 \le \|(x^*, \hat{y})\|_2$ and $\|(\hat{x}, \hat{y})\|_2 \le \|(\hat{x}, y^*)\|_2$ for all primal and dual solutions x^* and y^*. The desired result follows.) $\square$

5. Possibly unsolvable linear programs. We now consider a general linear program (1.1) which is possibly unsolvable. Let $\mathcal{R} : R_+^{n+m} \to R_+$ be a measure function for solvability of the problem (1.1), that is,

$$\mathcal{R}(x, y) = \|(A^T y - c)_+\|_1 + \|[-(Ax - b)]_+\|_1 + (c^T x - b^T y)_+.$$

Clearly, the value of the above function can also be viewed as a KKT residual corresponding to an approximate solution (x, y) of the linear program. Notice that $(x^*, y^*) \ge 0$ is a solution to the primal and the dual (1.1) and (1.2) if and only if $\mathcal{R}(x^*, y^*) = 0$. Thus a linear program is equivalent to the following global minimization problem:

$$\min\{\mathcal{R}(x, y) : (x, y) \ge 0\}. \tag{5.1}$$

We may refer (5.1) to the problem of minimizing the 1-norm solvability of linear program (1.1). By a basic idea of Mangasarian [16], the above problem can be reformulated as a linear programming problem. Indeed, by introducing nonnegative variables $(s, z, t) \in R_+^{n+m+1}$, the problem (5.1) can be equivalently transformed into the following linear program:

$$\begin{array}{ll} \min & e^T(s, z, t) \\ \text{s. t.} & A^T y - c \le s, \ -(Ax - b) \le z, \ c^T x - b^T y \le t, \\ & (x, y, s, z, t) \ge 0, \end{array} \tag{5.2}$$

where $e \in R^{n+m+1}$. This problem is always feasible. In fact, for any fixed $(x^0, y^0) \geq 0$, the vector $(x^0, y^0, s^0, z^0, t^0) \geq 0$ is feasible provided that $(s^0, z^0, t^0) > 0$ is sufficiently large. Since the objective function is nonnegative, the above linear program is always solvable, and hence the problem (5.1) has a global optimal solution. Let

$$c' := (0, 0, e) \in R^n \times R^m \times R^{n+m+1},\ b' := (-c, b, 0) \in R^n \times R^m \times R,$$

$$A' := \begin{bmatrix} O & -A^T & O & I & O \\ A & O & I & O & O \\ -c^T & b^T & O & O & 1 \end{bmatrix}_{(n+m+1)\times 2(m+n)+1},$$

and $u = (x, y, z, s, t)$. Then, (5.2) can be written as

$$\min\{(c')^T u : A'u \geq b', u \geq 0\}.$$

Replacing (c, A, b) by (c', A', b') and applying Algorithm 3.1 to the above problem, we can obtain the unique least 2-norm solution $(x^*, y^*, s^*, z^*, t^*)$ of the problem (5.2). We note that for any solution $(\hat{x}, \hat{y}, \hat{s}, \hat{z}, \hat{t})$ of (5.2), the following holds

$$\hat{z} = [-(A\hat{x} - b)]_+,\ \hat{s} = (A^T\hat{y} - c)_+,\ \hat{t} = (c^T\hat{x} - b^T\hat{y})_+.$$

Thus, for any solution $(\hat{x}, \hat{y}, \hat{s}, \hat{z}, \hat{t})$ of (5.2) we have

$$\begin{aligned} &\|(x^*,\ y^*,\ [-(Ax^* - b)]_+,\ (A^Ty^* - c)_+,\ (c^Tx^* - b^Ty^*)_+)\|_2 \\ &\leq \|(\hat{x},\ \hat{y},\ [-(A\hat{x} - b)]_+,\ (A^T\hat{y} - c)_+,\ (c^T\hat{x} - b^T\hat{y})_+)\|_2. \end{aligned} \tag{5.3}$$

When the linear program (1.1) or (1.2) is solvable, it is easy to see that any solution $(\hat{x}, \hat{y}, \hat{s}, \hat{z}, \hat{t})$ of (5.2) must satisfy that $\hat{s} = 0, \hat{z} = 0$ and $\hat{t} = 0$, and that $(\hat{x}, \hat{y})$ is a solution pair of the primal (1.1) and the dual (1.2). Conversely, if (x, y) is a solution pair to the primal and the dual, then $(x, y, 0, 0, 0)$ must be an optimal solution of the problem (5.2). Thus, for solvable linear program (1.1), the inequality (5.3) reduces to

$$\|(x^*, y^*, 0, 0, 0)\|_2 \leq \|(\hat{x}, \hat{y}, 0, 0, 0)\|_2$$

for all solution $(\hat{x}, \hat{y}, 0, 0, 0)$ of (5.2), which implies that x^* and y^* are the least 2-norm solutions of the primal (1.1) and the dual (1.2), respectively.

In summary, when applied to the linear program (5.2), Algorithm 3.1 is convergent no matter whether (1.1) is solvable or not. For solvable problems, the algorithm will converges to $(x^*, y^*, 0, 0, 0)$ where x^* and y^* are the least 2-norm solutions of the primal and the dual problems (1.1) and (1.2); otherwise, Algorithm 3.1 converges to a point which gives a minimal KKT residual.

6. Numerical results. While the linear system (3.2) is $(m + n)$-dimensional, we now point out that this system can be further reduced so that at each step, only an m- or n-dimensional linear system needs to be solved. In fact, (3.2) can be written as

$$(S^k + (\mu^k)^p X^k)\Delta x - X^k A^T \Delta y = \mu^k e - X^k \left((\mu^k)^p x^k - A^T y^k + c\right), \tag{6.1}$$

$$(Z^k + (\mu^k)^p Y^k)\Delta y + Y^k A \Delta x = \mu^k e - Y^k \left(Ax^k + (\mu^k)^p y^k - b\right). \tag{6.2}$$

When $m \geq n$, eliminating Δy leads to

$$\begin{aligned} M_k \Delta x = &\mu^k e - X^k \left((\mu^k)^p x^k - A^T y^k + c\right) \\ &+ X^k A^T (Z^k + (\mu^k)^p Y^k)^{-1} [\mu^k e - Y^k \left(A x^k + (\mu^k)^p y^k - b\right)], \end{aligned}$$

where M_k is an $n \times n$ matrix given by

$$M_k = S^k + (\mu^k)^p X^k + X^k A^T (Z^k + (\mu^k)^p Y^k)^{-1} Y^k A.$$

Thus, we can obtain Δx by solving the above system, and then set

$$\Delta y = (Z^k + (\mu^k)^p Y^k)^{-1} [\mu^k e - Y^k \left(A x^k + (\mu^k)^p y^k - b\right) - Y^k A \Delta x].$$

If $m \leq n$, by the same way, eliminating Δx from (6.1) and (6.2) yields

$$\begin{aligned} H_k \Delta y = &\mu^k e - Y^k \left((\mu^k)^p y^k + A x^k - b\right) \\ &- Y^k A (S^k + (\mu^k)^p X^k)^{-1} [\mu^k e - X^k \left((\mu^k)^p x^k - A^T y^k + c\right)], \end{aligned}$$

where H_k is an $m \times m$ matrix given by

$$H_k = Z^k + (\mu^k)^p Y^k + Y^k A (S^k + (\mu^k)^p X^k)^{-1} X^k A^T.$$

Since the system (3.2) has a unique solution, it follows that both M_k and H_k are nonsingular. Thus, at each step of Algorithm 3.1, we only need to factorize a matrix of size $\min(m, n) \times \min(m, n)$.

In numerical experiments, we took common parameters and starting points for all the test problems. Parameters were set as $p = 0.99, \sigma = 1e - 5, \alpha_1 = 0.9$ and $\alpha_2 = 0.85$. The starting point (x^0, y^0, s^0, z^0) was set as in Remark 3.1, where μ^0 and β were given by

$$\mu^0 = \max\left\{1, \left\| \begin{pmatrix} A^T y^0 - c \\ -A x^0 + b \end{pmatrix} \right\|_\infty \right\} + 1, \ \beta = (\eta + 1)/2.$$

Before stating our numerical results on some test problems, let us first see a very simple example with multiple solutions. Consider the following problem:

$$\min\{-x_1 - 2x_2 : x_1 + 2x_2 \leq 8, \ x_2 \leq 2, \ x_1, x_2 \geq 0\}.$$

It is easy to check that the solution set is $\{(x_1^*, x_2^*) = (4 + 4t, 2 - 2t) : 0 \leq t \leq 1\}$. Under a stopping rule of $\mu^k < 10^{-12}$, the following primal and dual solutions were obtained by the proposed algorithm,

$$x^* = (4.0000047139881082, \ 1.9999976430060873),$$

$$y^* = (0.9999999999886861, \ 1.6971276334429352e - 11)$$

with ∞-norm residual $\|\mathcal{F}_0(x^k, y^k, s^k, z^k)\|_\infty = 1.1314044977885658e{-}11$. Corresponding objective value of the original problem is -8.0000000000002828. We note that (4, 2) and (1, 0) are exact least 2-norm solutions of the primal and the dual problems, respectively, and the exact optimal objective value is -8. This example shows that the proposed algorithm does locate the least 2-norm solution of the problem.

TABLE 6.1

Name	Rows	Cols	Nonz	μ^k	$\|\mathcal{F}_0\|_\infty$	Objective values	CPU (secs)
beale	3	4	9	7.5e-11	2.0e-10	-1.25	.001
padberg	4	6	22	7.3e-09	1.8e-08	3.544147e-08	.001
refinery	22	14	63	9.6e-11	1.9e-09	-5.166833e+01	.08
william1	5	12	25	9.2e-11	1.6e-10	1.158471e-09	.002
william2	9	7	18	9.9e-11	3.3e-09	2.599999e+01	.05
william3	11	12	36	9.9e-10	1.3e-07	2.591899e+04	.4
afiro	35	32	117	9.9e-09	6.0e-06	-4.647531e+02	5.916
sc50a	70	48	182	8.9e-09	3.6e-06	-6.457507e+01	8.983
sc50b	70	48	170	9.9e-09	3.9e-06	-6.999999e+01	9.863
blend	117	83	789	9.9e-09	1.0e-06	-3.081215e+01	30.516
share2b	109	79	778	9.9e-09	3.7e-06	-4.157338e+02	37.600
sc105	150	103	402	9.8e-09	8.5e-06	-5.220206e+01	84.432
sc205	296	203	800	9.9e-09	2.8e-06	-5.220206e+01	325.632
scorpion	668	358	2526	9.8e-07	5.4e-05	1.878440e+03	> 500

We now give out a set of test examples and corresponding numerical results. We used $\mu^k \leq 10^{-10}$ or 10^{-8} as the stopping criterion for most of these test problems. All tests were carried out on a DEC Alpha V 4.0 machine. Results for 14 test problems were summarized in Table 6.1 and Table 6.2. The first problem was the well-known Beale's example and the second was Padberg's example ([20], pp.60). Both problems are cycling for simplex methods. The problem "refinery" can be found in [20]. The problems "william1", "william2", and "william3" were taken from [31] ($M = 50000$ was used in the problem "william3"). All other test problems here were taken from the collection of LP Data in NETLIB. In our code, all problems were transformed into the form of (1.1). To this end, all original inequalities "$\leq$" became "$\geq$" by multiplying both sides of inequalities by -1, and all equations were written equivalently as two inequalities. This preprocessing makes no change on the number of columns and keeps the sparsity of coefficient matrix A. However, the number of rows will be increased when the problem has equation constraints. The numbers of rows and nonzero entries of A in Table 6.1 are the resultant ones after this preprocessing. Under our stopping criterion, the computational optimal objective values, the values of μ^k, ∞-norm residual $\|\mathcal{F}_0(x^k, y^k, s^k, z^k)\|_\infty$, and CPU time were listed in Table 6.1. The computational primal and dual least 2-norm solutions for these test problems were given in Table 6.2, where only the first six components were listed due to the space limitation.

From our results, we note that Algorithm 3.1, using the initial strategy in Remark 3.1, is efficient for small-scale linear programs. However, the convergence rate of the algorithm becomes slow as the dimension of problems increases. The main reason might be that the stepsize of Armijo-type linear search may become smaller and smaller when iterates approach the least 2-norm solution. We also note that the matrices M_k and H_k are dense in general cases. Thus, when m and n are large, at each iteration a large and dense matrix needs to be factorized, which takes certain amount of CPU time. Thus, the current version of the algorithm is not so efficient for solving large-scale problems. Some modified versions of the algorithm are worth studying in the future in order to improve its convergence rate. A possible method is to use certain approximate Newton step to accelerate the iteration as we have done for nonlinear complementarity problems in [37].

Table 6.2

Name	Primal and dual least 2-norm solution (x^*, y^*)
beale	$x^* = (1.00000, 0.00000, 1.00000, 0.00000)$ $y^* = (0.00000, 1.50000, 1.25000)$
padberg	$x^* = (1.29177, 0.00000, 0.64588, 0.00000, 0.64588, 0.00000)$ $y^* = (0.91360, 0.91342, 0.91360, 0.91342)$
refinery	$x^* = (15.00000, 10.00000, 3.50000, 6.25000, 8.00000, 1.55000, ...)$ $y^* = (0.06333, 1.62166, 0.81166, 4.99246, 4.99246, 2.51578, ...)$
william1	$x^* = (0.00000, 0.00000, 0.00000, 0.00000, 0.00000, 0.00000, ...)$ $y^* = (0.00000, 0.00000, 0.00000, 0.00000, 0.00000)$
william2	$x^* = (0.00000, 4.16154, 17.00000, 7.00000, 17.00000, 22.00000, ...)$ $y^* = (0.00000, 0.00000, 1.00000, 0.00000, 1.00000, 1.00000, ...)$
william3	$x^* = (0.00000, 0.00000, 39.00000, 87.00000, 56.00000, 0.00000, ...)$ $y^* = (0.00000, 26.00000, 5.00000, 111.00732, 94.00864, 97.00838, ...)$
afiro	$x^* = (80.00000, 25.50000, 54.50000, 84.79999, 36.85030, 0.00000, ...)$ $y^* = (0.65036, 0.91201, 0.34477, 0.22857, 0.91201, 0.91201, ...)$
sc50a	$x^* = (0.00000, 16.56869, 64.57507, 64.57507, 64.57507, 0.00000, ...)$ $y^* = (0.00000, 0.13869, 0.91201, 0.81390, 0.85202, 0.78381, ...)$
sc50b	$x^* = (29.99999, 28.00000, 42.0000, 69.99999, 69.99999, 29.99999, ...)$ $y^* = (0.05836, 0.91201, 0.91201, 0.82870, 0.82871, 0.82871, ...)$
blend	$x^* = (20.94480, 10.17092, 11.24735, 2.98109, 0.65970, 0.47592, ...)$ $y^* = (0.21613, 0.22386, 0.26003, 0.26003, 0.25294, 0.25983, ...)$
share2b	$x^* = (1.95814, 2.02325, 0.00000, 0.00000, 0.00000, 0.00000, ...)$ $y^* = (0.12564, 0.00000, 0.00000, 0.00000, 0.00000, 0.33250, ...)$
sc105	$x^* = (0.00000, 10.84845, 52.20206, 52.20206, 52.20206, 0.00000, ...)$ $y^* = (0.00000, 0.16419, 0.91201, 0.79709, 0.84241, 0.76248, ...)$
sc205	$x^* = (0.00000, 10.84845, 52.20206, 52.20206, 52.20206, 0.00000, ...)$ $y^* = (0.00000, 0.16136, 0.91201, 0.79872, 0.84356, 0.76481, ...)$
scorpion	$x^* = (0.00871, 0.00211, 0.00023, 0.00452, 1.42494, 0.00250, ...)$ $y^* = (0.9293, 113.1449, 115.4149, 115.4149, 0.0000, 421.2979, ...)$

7. Conclusions. In this paper, we have introduced a new concept of regularized central path for linear programs, which is different from the conventional central path. The regularized central path always exists for all linear programs, even if the linear program is unsolvable. If a linear program is solvable, the regularized central path converges, as the parameter μ tends to zero, to the unique least 2-norm solution of the linear program. As a result, we propose in this paper a regularized central path-based path-following algorithm to solve linear programming problems. This is a new alternative algorithm for locating the least 2-norm solution of a linear program. When applied to the equivalent problem (5.2), the iterative sequence generated by this algorithm is always convergent no matter whether the problem is solvable or not. If the primal problem is solvable, the limiting point of the sequence is the least 2-norm solution; otherwise, the limiting point gives a minimal KKT residual.

It should be pointed out that most of the existing algorithms for the least-norm solution of the linear program are akin to the canonical Tikhonov regularization method.

One significance of the proposed algorithm is that it introduces the framework of interior-point methods into the canonical Tikhonov regularization method. As a result, the proposed algorithm can be viewed as a new effective implementation version of the classical Tikhonov regularization method. In addition, the convergence of the algorithm needs no assumption when applied to the reformulated problem (5.2).

From our results, some interesting problems arise: What is the rate of convergence of Algorithm 3.1? Can certain modified version of the algorithm be superlinearly (or quadratically) convergent in the neighborhood of the least 2-norm solution of a linear program? Can the least 2-norm solution of a linear program be solved in polynomial time? We believe that these problems are worth studying in the future.

Acknowlegements. We thank two anonymous referees for their helpful comments and suggestions. We also thank Professor Z.L. Wei for helping on the code.

REFERENCES

[1] M. J. Best, *Equivalence of some quadratic programming algorithms,* Math. Programming, 30 (1984), pp. 71–87.

[2] F. Facchinei, *Structural and stability properties of P_0 nonlinear complementarity problems*, Math. Oper. Res., 23 (1998), pp. 735–749.

[3] F. Facchinei and C. Kanzow, *Beyond monotonicity in regularization methods for nonlinear complementarity problems*, SIAM J. Control Optim., 37 (1999), pp. 1150–1161.

[4] G. Isac, *Tikhonov's regularization and the complementarity problem in Hilbert space,* J. Math. Anal. Appl., 174 (1991), pp. 53–66.

[5] C. Kanzow, H. Qi and L. Qi, *On the minimum norm solution of linear programming,* Department of Mathematics, Center for Optimization and Approximation, Hamburg, Germany, 2000.

[6] K. C. Kiwiel, *Finding normal solutions in piecewise linear programming,* Appl. Math. Optim., 32 (1995), pp. 235–254.

[7] K. C. Kiwiel, *Iterative schemes for the least 2-norm solution of piecewise linear programs,* Linear Algebra Appl., 229 (1995), pp. 1–7.

[8] K. C. Kiwiel, *A dual method for certain positive semidefinite quadratic programming problems,* SIAM J. Sci. Statist. Comput. 10 (1989), pp. 175–186.

[9] M. Kojima, N. Megiddo, and T. Noma, *Homotopy continuation methods for nonlinear complementarity problems,* Math. Oper. Res., 16 (1991), pp. 754–774.

[10] Y. Y. Lin and J. S. Pang, *Iterative methods for large convex quadratic programs: A survey,* SIAM J. Control Optim., 25 (1987), pp. 383–411.

[11] S. Lucidi, *A new result in the theory and computation of the least-norm solution of a linear program*, J. Optim. Theory Appl., 55 (1987), pp. 103–117.

[12] S. Lucidi, *A finite algorithm for the least two-norm solution of a linear program,* Optimization, 18 (1987), pp. 809–823.

[13] O. L. Mangasarian, *Uniqueness of solution in linear programming,* Linear Algebra and Appl., 25 (1979), pp. 151–162.

[14] O. L. Mangasarian, *Iterative solution of linear programs,* SIAM J. Numer. Anal., 18 (1981), pp. 606–614.

[15] O. L. Mangasarian, *Least-norm linear programming solution as an unconstrained minimization problem,* J. Math. Anal. Appl., 92 (1983), pp. 240–251.

[16] O. L. Mangasarian, *Normal solutions of linear programs,* Math. Programming Study, 22 (1984), pp. 206–216.

[17] O. L. Mangasarian, *Least norm solution of monotone linear complementarity problems,* Functional Analysis, Optimization, and Mathematical Economics, pp. 217–221, Oxford Univ. Press, New York, 1990.

[18] O. L. Mangasarian and R. De Leone, *Error bounds for strongly convex programs and (super)linearly convergent iterative schemes for the least 2-norm solution of linear programs,* Appl. Math. Optim., 17 (1988), pp. 1–14.

[19] O. L. Mangasarian and R. R. Meyer, *Nonlinear perturbation of linear programs,* SIAM J. Control and Optim., 17 (1979), pp. 745–757.

[20] M. Padberg, *Linear Optimization and Extensions,* Springer-Verlag, Berlin, 1995.

[21] T. D. Parsons and A. W. Tucker, *Hybrid program: linear and least-distance,* Math. Programming, 1 (1971), pp. 153–167.

[22] S. M. Robinson, *A characterization of stability in linear programming,* Oper. Res., 25 (1977), pp. 435–447.

[23] R. Saigal, *Linear Programming: A Modern Integrated Analysis,* Kluwer Academic Publishers, 1995.

[24] P. W. Smith and H. Wolkowicz, *A nonlinear equation for linear programming,* Math. Programming, 34 (1986), pp. 235–238.

[25] V. D. Skarin, *Methods for the correction of ill-posed problems of linear and convex programming by using a sequential programming approach,* In: Parametric Optimization and Ill-Posed Problems in Mathematical Optimization, pp. 130–144, Seminarberichte, 81, Humboldt Univ., Berlin, 1986.

[26] P. K. Subramanian, *A note on least two norm solutions of monotone complementarity problems,* Appl. Math. Lett., 1 (1988), pp. 395–397.

[27] R. Sznajder and M. S. Gowda, *On the limiting behavior of the trajectory of regularized solutions of P_0 complementarity problems,* Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, M. Fukushima and L. Qi (Eds), Kluwer Academic Publishers, pp. 371–379, 1998.

[28] A. W. Tucker, *A least-distance approach to quadratic programming,* In: Mathematics of the Decision Sciences, Part I, Vol. 11 of Lecture in Applied Mathematics, pp. 163–176, (Eds. G.B. Dantzig and A.F. Veinott), 1968.

[29] A. W. Tucker, *Least-distance programming*, In: Proceedings of the Princeton Symposium on Mathematical Programming, pp. 583–588, (Ed. H.W. Kuhn), Princeton University Press, Princeton, N.J., 1971.

[30] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-posed Problems,* Halsted Press, Wiley, New York, 1977.

[31] H. P. Williams, *Model Building in Mathematical Programming*, John Wiley & Sons, 1995.

[32] P. Wolfe, *Algorithm for a least-distance programming problem,* Math. Prog. Study, 1 (1974), pp. 190–205.

[33] P. Wolfe, *Finding the nearest point in a polytope,* Math. Programming, 11 (1976), pp. 128–149.

[34] Y. Ye, *Interior-Point Algorithms: Theory and Analysis,* John Wiley and Sons, Chichester, 1997.

[35] Y. B. Zhao and D. Li, *On a new homotopy continuation trajectory for complementarity problems,* Math. Oper. Res., 26 (2001), pp. 119-146.

[36] Y. B. Zhao and D. Li, *Existence and limiting behavior of non-interior-point trajectory for nonlinear complementarity problems without strict feasibility condition,* SIAM J. Control Optim., 40 (2001), pp. 898-924.

[37] Y. B. Zhao and D. Li, *A new path-following algorithm for nonlinear P_* complementarity problems*, Department of SEEM, Chinese University of Hong Kong, Hong Kong, November, 2000.