On the Log-exponential Trajectory of Linear Programming

Jie Sun*
Singapore-MIT Alliance and Department of Decision Sciences
National University of Singapore, Republic of Singapore 119260

Liwei Zhang†
CORA, Department of Applied Mathematics,
Dalian University of Technology, Dalian 116024, China
Center for Advanced Information Systems, School of Computer Engineering
Nanyang Technological University, Republic of Singapore 639798

Abstract. Development in interior point methods has suggested various solution trajectories, also called central paths, for linear programming. In this paper we define a new central path through a log-exponential perturbation to the complementarity equation in the Karush-Kuhn-Tucker system. The behavior of this central path is investigated and an algorithm is proposed. The algorithm can compute an ϵ-optimal solution at a superlinear rate of convergence.

Keywords: damped Newton method, interior point method, linear programming, log-exponential function, superlinear convergence

1. Introduction

The development of interior point methods for linear programming has caught great attention among researchers in optimization in the last two decades. Interior point methods are not only computationally competitive, but also motivational in many new directions in theoretical research. For instance, the notion of central path plays a significant role in interior point methods for linear programming (LP) (e.g. see Ye (1997)). Let

$$\mathcal{P} = \min\{c^T x \mid Ax = b, x \geq 0\}, \quad \mathcal{D} = \{\max b^T y \mid s = c - A^T y, s \geq 0\}$$

be the pair of standard linear programming problems, where $b, y \in \mathbf{R}^m, c, x, s \in \mathbf{R}^n, A \in \mathbf{R}^{m\times n}$, and "$T$" represents the transpose. The central path refers to the primal-dual solution set $\{(x(t), y(t), s(t)) \mid$

* This author's work was partially supported by Singapore-MIT Alliance and Grants R314-000-026/028/042-112 of National University of Singapore. Email:jsun@nus.edu.sg. Fax: (65)6779-2621
† This author's work was supported by the Youth NSF of China under No.10001007 and High Performance Container Operation Founded by NSTB, MPA and NTU. Email: aslwzhang@ntu.edu.sg

$t > 0\}$ to the following Karush-Kuhn-Tucker (KKT) system

$$\begin{cases} Ax = b \\ A^T y + s = c \\ x \circ s = te, x > 0, s > 0 \end{cases},$$

where t is a parameter, e is the vector of ones and $x \circ s$ is the vector Hadamard product: $x \circ s = [x_i s_i]_1^n$. Mainstream interior point methods such as the primal-dual path-following methods find a sequence of approximate solutions to the above system as $t \downarrow 0$, starting from some approximate solution with $t = t_0$.

In this paper we consider a different central path and study its mathematical properties. A new method for LP is proposed based on this central path. Rather than the standard pair of LP, the new central path is based on the Karmarkar pair of linear programming. Namely, let

$$\mathcal{F}'_P(\mu) = \{x' \in \mathbf{R}^{n+1} \,|\, [A, -b]x' = 0,\ e_{[n+1]}^T x' = 1, x' \geq 0\},$$

where $x' \in \mathbf{R}^{n+1}$ and $e_{[n+1]}$ is the vector of ones in $\mathbf{R}^{n+1}$. Then the primal problem is

$$\min \left\{ c^T x'_{[n]} - \mu x'_{n+1} \,|\, x' \in \mathcal{F}'_P(\mu) \right\}. \qquad (\text{ALP}(\mu))$$

The dual of (ALP)(μ) is the following problem

$$\max \left\{ y'_{m+1} \,|\, (y', s') \in \mathcal{F}'_D(\mu) \right\}, \qquad (\text{ALD}(\mu))$$

where

$$\mathcal{F}'_D(\mu) = \left\{ (y', s') \in \mathbf{R}^{m+1} \times \mathbf{R}^{n+1} \,|\, \begin{array}{r} A^T y'_{[m]} + y'_{m+1} e + s_{[n]} = c, \\ b^T y'_{[m]} + y'_{m+1} + s_{n+1} = \mu, s \geq 0 \end{array} \right\}$$

and

$$y' = \begin{bmatrix} y'_{[m]} \\ y'_{m+1} \end{bmatrix},\ y'_{[m]} \in \mathbf{R}^m \text{ and } y'_{m+1} \in \mathbf{R}.$$

It should be noted that the dual problem (ALD)(μ) can also be written as the following min-max problem

$$\min \left\{ -y'_{m+1} \,|\, -y'_{m+1} = \max[A^T y'_{[m]} - c, b^T y'_{[m]} - \mu] \right\}. \qquad (1)$$

It is well known that the standard form and the Karmarkar form of LP are equivalent in the following sense.

Proposition 1. Assume that the feasible sets of $\mathcal{P}$ and $\mathcal{D}$ are bounded and both have nonempty relative interiors. If x^* is a solution to (LP), then $(x^*, 1)/(1 + e^T x^*)$ is a solution to (ALP)(μ^*), where μ^* is the optimal value of (LP). If $\hat{x}'$ is a solution to (ALP)(μ^*), then $\hat{x}'_{n+1} \neq 0$ and $\hat{x}'_{[n]}/\hat{x}'_{n+1}$ is a solution to (LP).

If μ^* is unknown, then an auxiliary procedure can be designed to find μ^* in practice. Since the detail is irrelevant to our analysis below, we simply assume $\mu = \mu^*$ in the sequel. We also make the blanket assumption that both $\mathcal{P}$ and $\mathcal{D}$ have bounded feasible sets with nonempty relative interiors in view of Proposition 1.

The solution set of (ALP)(μ) and (ALD)(μ), denoted by $\mathcal{F}^*(\mu)$, is defined by

$$\mathcal{F}^*(\mu) = \left\{(x', y', s') \in \mathcal{F}'_P(\mu) \times \mathcal{F}'_D(\mu) \,|\, s'_i x'_i = 0, i = 1, \ldots, n+1\right\}.$$

Let us define a new central path as

$$\mathcal{C}_t = \left\{(x', y', s') \in \mathcal{F}'_P(\mu) \times \mathcal{F}'_D(\mu) \,\middle|\, \begin{array}{c} t \log x'_i + s'_i = 0,\ x'_i > 0, \\ i = 1, \ldots, n+1, t > 0 \end{array}\right\},$$

or by using vector notation,

$$\mathcal{C}_t = \left\{(x', y', s') \in \mathcal{F}'_P(\mu) \times \mathcal{F}'_D(\mu) \mid s' = -t \log x', x' > 0, t > 0\right\}, \quad (2)$$

where $\log(x')$ denotes the vector of $\mathbf{R}^{n+1}$ whose i-th component is $\log x'_i$. Then for $(x', y', s') \in \mathcal{C}_t$ one has that

$$1 = \sum_{i=1}^{n+1} x'_i = \left[\sum_{i=1}^{n+1} \exp\left(t^{-1}(A_i^T y'_{[m]} - c_i)\right)\right] \exp(t^{-1} y'_{m+1}),$$

where $A_{n+1} \equiv -b$ and $c_{n+1} \equiv -\mu$. Therefore,

$$-y'_{m+1} = t \log \left[\sum_{i=1}^{n+1} \exp\left(t^{-1}(A_i^T y'_{[m]} - c_i)\right)\right], \quad (3)$$

and the term on the right-hand side in the above equality is just the log-exponential (log-exp for short) function of

$$F(y'_{[m]}, \mu) \equiv \begin{bmatrix} A^T y'_{[m]} - c \\ \mu - b^T y'_{[m]} \end{bmatrix}.$$

As $t \downarrow 0$ the log-exp function uniformly approaches the vector-max function, namely

$$0 \leq \text{logexp}(z) - \text{vecmax}(z) \leq t \log n,$$

where

$$\text{logexp}(z) = \log\left[\sum_{i=1}^{n} \exp(z_i)\right], \quad \text{vecmax}(z) = \max\{z_1, \ldots, z_n\}$$

for $z \in \mathbf{R}^n$ (see 1.30 of Rockafellar and Wets (1998)). Thus, the points on $\mathcal{C}_t$ can be interpreted as approximate optimal solutions of ALD(μ) according to (1).

In this paper we study the solution trajectory $\mathcal{C}_t$. The theory of this solution trajectory is established based on the properties of log-exp function, which were studied by many authors, see Rockafellar and Wets (1998), Templeman and Li (1987), Li (1991, 1992), and Peng and Lin (1999). Also, Chen and Mangasarian (1996) used the recession function of $\log\exp(z)$ when $z \in \mathbf{R}^2$ to solve complementarity problems.

In Section 2 we study the new solution trajectory and neighborhoods of the trajectory. In Section 3 we give an unconstrained approach to find a point on the solution trajectory. The results obtained are connected to Fang (1992). Some concluding remarks are made in Section 4.

2. The solution trajectory

Let

$$\Phi(y'_{[m]}, \mu) \equiv \text{vec}\max[F(y'_{[m]}, \mu)] = \text{vec}\max\begin{bmatrix} A^T y'_{[m]} - c \\ \mu - b^T y'_{[m]} \end{bmatrix}, \tag{4}$$

and

$$\Phi_t(y'_{[m]}, \mu) \equiv$$
$$t\log\left[\sum_{j=1}^{n} \exp\left(t^{-1}(A_j^T y'_{[m]} - c_j)\right) + \exp\left(t^{-1}(\mu - b^T y'_{[m]})\right)\right]. \tag{5}$$

Then $\Phi_t(y'_{[m]}, \mu)$ approximates $\Phi(y'_{[m]}, \mu)$ by the following inequalities

$$\Phi_t(y'_{[m]}, \mu) - t\log(n+1) \leq \Phi(y'_{[m]}, \mu) < \Phi_t(y'_{[m]}, \mu). \tag{6}$$

Remark 1. Note that both Φ and Φ_t are convex functions and for any β the level set of Φ_t defined by $\{y'_{[m]} \mid \Phi_t(y'_{[m]}, \mu) \leq \beta\}$ is contained in the corresponding level set of Φ due to (6). Form the blanket assumption we know that one of the level sets of Φ is bounded, hence one of the level sets of Φ_t is also bounded, which is equivalent to that all of the level sets of Φ_t are bounded by convex analysis.

Note that if $(x', y', s') \in \mathcal{C}_t$, then it follows from (6) that

$$-y'_{m+1} > F_i(y'_{[m]}, \mu), i = 1, \dots, n+1,$$

and therefore,

$$s'_i = -F_i(y'_{[m]}, \mu) - y'_{m+1} > 0, \; i = 1, \dots, n+1, \tag{7}$$

i.e., $(y', s') \in \mathrm{rint}\mathcal{F}'_D(\mu)$, where "rint" denotes the relative interior in the usual sense of convex analysis. Thus, $\mathcal{C}_t$ could be expressed in the form

$$\mathcal{C}_t = \{(x', y', s') \,|\, \tilde{A}x' = 0, e^T_{[n+1]}x' = 1, \log(x') + t^{-1}s' = 0, s' = \bar{c}(\mu) - \bar{A}^T y'\}.$$

where

$$\tilde{A} = (A, -b) \text{ and } \bar{c}(\mu) = (c^T, -\mu)^T. \tag{8}$$

Later in (21) we will show that (x', y', z') exists and is unique for any $t > 0$. In the remainder of this section, we will prove that $\mathcal{C}_t$ is a solution trajectory in the following sense

$$\lim_{t \downarrow 0} \mathrm{dist}((x'(t), y'(t), s'(t)), \mathcal{F}^*(\mu)) = 0,$$

for $(x'(t), y'(t), s'(t)) \in \mathcal{C}_t$. Our proof is based on minimizing $\Phi_t(y'_{[m]}, \mu)$. It should be pointed out that $\Phi_t(y'_{[m]}, \mu)$ is continuously differentiable which overcomes the nonsmoothness of $\Phi(y'_{[m]}, \mu)$. We introduce some notations as follows, which will be used in the sequel:

$$\xi'_i(y'_{[m]}, t) = \exp\left[t^{-1}\left(A_i^T y'_{[m]} - c_i - \Phi_t(y'_{[m]}, \mu)\right)\right], \; i = 1, \dots, n,$$

$$\xi'_{n+1}(y'_{[m]}, t) = \exp\left[t^{-1}\left(\mu - b^T y'_{[m]} - \Phi_t(y'_{[m]}, \mu)\right)\right],$$

$$E(y'_{[m]}, t) = \mathrm{diag}_{1 \le i \le n+1}(\xi'(y_{[m]}, t)),$$

$$H(y'_{[m]}, t) = E(y'_{[m]}, t) - E(y'_{[m]}, t)e_{[n+1]}e^T_{[n+1]}E(y'_{[m]}, t).$$

It is easy to verify the following properties:

(P1) $\sum_{i=1}^{n+1} \xi_i(y'_{[m]}, t) = 1$;
(P2) $E(y'_{[m]}, t)$ is positive definite;
(P3) $H(y'_{[m]}, t) = H(y'_{[m]}, t)E(y'_{[m]}, t)^{-1}H(y'_{[m]}, t)$;
(P4) Null($H(y'_{[m]}, t)$)= Span ($e_{[n+1]}$);

(P5) $\nabla(\Phi_t(y'_{[m]},\mu))$ and $\nabla^2(\Phi_t(y'_{[m]},\mu))$ have following expressions

$$\begin{aligned}\nabla(\Phi_t(y'_{[m]},\mu)) &= \textstyle\sum_{j=1}^{n+1} \xi'_j(y'_{[m]},t)A_j \\ &= \tilde{A}E(y'_{[m]},t)e_{[n+1]},\end{aligned} \tag{9}$$

and

$$\begin{aligned}&\nabla^2(\Phi_t(y'_{[m]},\mu)) \\ = \;& t^{-1}\left\{\sum_{j=1}^{n+1}\xi'_j(y'_{[m]},t)A_jA_j^T - (\sum_{j=1}^{n+1}\xi'_j(y'_{[m]},t)A_j)(\sum_{j=1}^{n+1}\xi'_j(y'_{[m]},t)A_j)^T\right\} \\ = \;& t^{-1}\left\{\tilde{A}(E(y'_{[m]},t) - E(y'_{[m]},t)e_{[n+1]}e_{[n+1]}^TE(y'_{[m]},t))[A,-b]^T\right\} \\ = \;& t^{-1}[A,-b]H(y'_{[m]},t)\tilde{A}^T,\end{aligned}$$

where $A_{n+1} = -b$.

Lemma 1. If A is of full row rank, then $\Phi_t(y'_{[m]},\mu)$ is strictly convex in $y'_{[m]}$.

Proof. From (P5) and (P3), we can write $\nabla^2(\Phi_t(y'_{[m]},\mu))$ as

$$\nabla^2(\Phi_t(y'_{[m]},\mu)) = t^{-1}\tilde{A}H(y'_{[m]},t)E(y'_{[m]},t)^{-1}H(y'_{[m]},t)\tilde{A}^T, \tag{10}$$

which implies that $\nabla^2(\Phi_t(y'_{[m]},\mu))$ is positive semi-definite. Now, we prove that it is in fact a positive definite matrix. For

$$H(y'_{[m]},t)\tilde{A}^Tz = 0,$$

it follows from (P4) that

$$\tilde{A}^Tz = (A,-b)^Tz \in \mathrm{Span}\{e_{[n+1]}\}.$$

If $(A,-b)^Tz = \lambda e_{[n+1]}$ and $\lambda \neq 0$, then $\mathcal{F}'_P(\mu) = \emptyset$, which is a contradiction. Therefore $[A,-b]^Tz = 0$, which implies $z = 0$ since A is of full row rank. Thus the matrix

$$H(y'_{[m]},t)(A,-b)^T$$

is of full column rank, $\nabla^2(\Phi_t(y'_{[m]},\mu))$ is positive definite, and $\Phi_t(y'_{[m]},\mu)$ is strictly convex in $y'_{[m]}$. □

Since $\Phi_t(y'_{[m]}, \mu)$ is strictly convex in $y'_{[m]}$ and has bounded level sets, it has a unique minimizer, denoted here by $y'_{[m]}(t, \mu)$, namely

$$y'_{[m]}(t, \mu) \equiv \operatorname{argmin}\{\Phi_t(y'_{[m]}, \mu) \,|\, y'_{[m]} \in \mathbf{R}^m\}.$$

For convenience, we denote $\xi'(t, \mu) \equiv \xi'(y'_{[m]}(t, \mu), t)$ Then

$$\nabla_{y'_{[m]}} \Phi_t(y'_{[m]}(t, \mu), \mu) = 0,$$

or

$$\sum_{j=1}^{n+1} \xi'_j(t, \mu) A_j = 0, \tag{11}$$

namely

$$\tilde{A} E(y'_{[m]}(t, \mu), t) e_{[n+1]} = 0. \tag{12}$$

From (11) (or (12)), (P1) and the positiveness of $\xi'(t, \mu)$, we obtain the following important inclusion relationship

$$\xi'(t, \mu) \in \operatorname{rint} \mathcal{F}'_P(\mu). \tag{13}$$

The following theorem shows that $\xi'(t, \mu)$ is an approximate (interior) solution to (ALP)(μ) and $(y'(t, \mu), -\Phi_t(y'(t, \mu), \mu))$ is an approximate (interior) solution to (ALD)(μ).

Theorem 1. Let $y'^*_{m+1}(\mu)$ be the optimal value of (ALD)(μ), then

$$\begin{aligned}
(y'_{[m]}(t, \mu), -\Phi_t(y'_{[m]}(t, \mu), \mu)) \quad & \in \operatorname{rint}\mathcal{F}'_D(\mu), & (14)\\
\xi'(t, \mu) \quad & \in \operatorname{rint}\mathcal{F}'_P(\mu), & (15)\\
-t \log(n+1) + c^T \xi'_{[n]}(t, \mu) - \mu \xi'_{n+1}(t, \mu) \quad & \le y'^*_{m+1}(\mu) \le \\
\le c^T \xi'_{[n]}(t, \mu) - \mu \xi'_{n+1}(t, \mu), & & (16)
\end{aligned}$$

and

$$-\Phi_t(y'_{[m]}(t, \mu), \mu) \le y'^*_{m+1}(\mu) \le -\Phi_t(y'_{[m]}(t, \mu), \mu) + t \log(n+1). \tag{17}$$

Proof. From the definition of $\xi'(t, \mu)$, we have

$$\begin{aligned}
&t \log(\xi'_i(t, \mu)) = A_i^T y'_{[m]}(t, \mu) - c_i - \Phi_t(y'_{[m]}(t, \mu), \mu), \; i = 1, \ldots, n,\\
&t \log(\xi'_{n+1}(t, \mu)) = \mu - b^T y'_{[m]}(t, \mu) - \Phi_t(y'_{[m]}(t, \mu), \mu),
\end{aligned}$$

or in a compact form

$$\begin{aligned}
&A^T y'_{[m]}(t, \mu) - c - \Phi_t(y'_{[m]}(t, \mu), \mu) e_{[n]} = t \log(\xi'_{[n]}(t, \mu)),\\
&\mu - b^T y'_{[m]}(t, \mu) - \Phi_t(y'_{[m]}(t, \mu), \mu) = t \log(\xi'_{n+1}(t, \mu)),
\end{aligned} \tag{18}$$

where $\log(\xi'_{[n]}(t,\mu)) \in \mathbf{R}^n$ denotes the vector whose i-th component is $\xi'_i(t,\mu)$ for $i = 1,\ldots,n$. In view of the definition of $\xi'(t,\mu)$, one has that

$$\begin{array}{l} A^T y'_{[m]}(t,\mu) - \Phi_t(y'_{[m]}(t,\mu),\mu)e_{[n]} < c, \\ -b^T y'_{[m]}(t,\mu) - \Phi_t(y'_{[m]}(t,\mu),\mu) < -\mu, \end{array}$$

which implies the validity of (14). The inclusion (15) comes from (13). Expressions in (18) may be re-written as

$$\tilde{A}^T y'_{[m]}(t,\mu) - \Phi_t(y'_{[m]}(t,\mu),\mu)e_{[n+1]} = \bar{c}(\mu) + t\log(\xi'(t,\mu)).$$

Premultiplying the above equality by $\xi'(t,\mu)^T$, we have

$$-\Phi_t(y'_{[m]}(t,\mu),\mu) = c^T\xi'_{[n]}(t,\mu) - \mu\xi'_{n+1}(t,\mu) + tH_{n+1}(\xi'(t,\mu)),$$

where

$$H_{n+1}(\xi'(t,\mu)) = \sum_{j=1}^{n+1} \xi'_j(t,\mu)\log\xi'_j(t,\mu)$$

is the negative entropy with respect to $\xi'(t,\mu)$. It is easy to verify that

$$\min\{H_{n+1}(\xi') \mid \sum_{j=1}^{n+1}\xi'_j = 1, \xi' \geq 0\} = -\log(n+1).$$

Thus we have

$$\begin{array}{rl} & -\Phi_t(y'_{[m]}(t,\mu),\mu) < c^T\xi'_{[n]}(t,\mu) - \mu\xi'_{n+1}(t,\mu) \\ \leq & -\Phi_t(y'_{[m]}(t,\mu),\mu) + t\log(n+1). \end{array} \tag{19}$$

Since $(y'_{[m]}(t,\mu), -\Phi_t(y'_{[m]}(t,\mu),\mu)) \in \operatorname{rint}\mathcal{F}'_D(\mu)$ and $\xi'(t,\mu) \in \operatorname{rint}\mathcal{F}'_P(\mu)$, it follows from the duality theory of linear programming that

$$-\Phi_t(y'_{[m]}(t,\mu),\mu) \leq y'^*_{m+1}(\mu) \leq c^T\xi'_{[n]}(t,\mu) - \mu\xi'_{n+1}(t,\mu). \tag{20}$$

Combining inequalities (19) and (20), we obtain inequalities (16) and (17). The proof is completed.
Since

$$\begin{array}{rcl} \log\xi'_i(t,\mu) & = & t^{-1}\{[A_i^T y'_{[m]}(t,\mu) - c_i] - \Phi_t(y'_{[m]}(t,\mu),\mu)\} \\ & = & -t^{-1}(c_i - A_i^T y'_{[m]}(t,\mu) - (-\Phi_t(y'_{[m]}(t,\mu),\mu))), \end{array}$$

or

$$\log\xi'(t,\mu) = -t^{-1}(\bar{c}(\mu) - \tilde{A}^T y'(t,\mu)), y'(t,\mu) = (y'_{[m]}(t,\mu), -\Phi_t(y'_{[m]}(t,\mu),\mu)),$$

and

$$\tilde{A}\xi'(t,\mu) = 0, e_{[n+1]}^T \xi'(t,\mu) = 1.$$

We obtain that

$$(\xi'(t,\mu), y'(t,\mu), s'(t,\mu)) \in \mathcal{C}_t$$

with $s'(t,\mu) = \bar{c}(\mu) - \tilde{A}^T y'(t,\mu)$. Noting $y'_{[m]}(t,\mu)$ is the unique minimizer of $\Phi_t(y_{[m]},\mu)$, we have that if a point $(x', y', s') \in \mathcal{C}_t$ corresponding to $t > 0$, i.e., it satisfies

$$\begin{array}{l} \tilde{A}x' = 0, \\ e_{[n+1]}^T x' = 1, \\ s' = \bar{c}(\mu) - \tilde{A}^T y', \\ \log x' + t^{-1} s' = 0, \end{array}$$

then it is just $(x'(t,\mu), y'(t,\mu), s'(t,\mu))$. Based on this fact, $\mathcal{C}_t$ may also be expressed as

$$\begin{aligned} \mathcal{C}_t = \{ & (\xi'(t,\mu), (y'_{[m]}(t,\mu), -\Phi_t(y'_{[m]}(t,\mu),\mu)), -t \log \xi'(t,\mu)) \\ & \mid y'_{[m]}(t,\mu) = \text{argmin } \Phi_t(y'_{[m]},\mu), t > 0\}. \end{aligned} \tag{21}$$

To illustrate that $\mathcal{C}_t$ defined by (2) or (21) is actually a solution trajectory to (ALP)(μ) and (ALD)(μ), we must show that the distance of $\xi'(t,\mu)$ ($y'(t,\mu)$) and the solution set of (ALP)(μ)((ALD)(μ)) tends to zero as t tends to zero.

Theorem 2. Let

$$(x'(t,\mu), y'(t,\mu), s'(t,\mu)) \equiv (\xi'(t,\mu), (y'_{[m]}(t,\mu), -\Phi_t(y'_{[m]}(t,\mu),\mu))$$

and $-t \log \xi'(t,\mu)) \in \mathcal{C}_t, t > 0$ Then

$$\lim_{t \downarrow 0} \text{dist}(x'(t,\mu), X'_\mu) = 0, \quad \lim_{t \downarrow 0} \text{dist}(y'(t,\mu), Y'_\mu) = 0, \tag{22}$$

where $X'(\mu)$ and $Y'(\mu)$ are optimal solution sets of (ALP)(μ) and (ALD)(μ), respectively.

Proof. Let

$$X'_\mu(t) \equiv \{x' \in \mathcal{F}'_P(\mu) \mid \bar{c}(\mu)x' \le -\Phi_t(y'_{[m]}(t,\mu),\mu) + t\log(n+1)\}$$

and

$$Y'_\mu(t) \equiv \{y' \mid (y', s') \in \mathcal{F}'_D(\mu), y'_{m+1} \ge -\Phi_t(y'_{[m]}(t,\mu),\mu)\}.$$

In view of (16) and (17), we have $X'_\mu \subset X'_\mu(t)$ and $Y'_\mu \subset Y'_\mu(t)$. Noting that $\Phi_t(y'_{[m]}, \mu)$ is monotonically increasing with respect to $t > 0$, we have that

$$\Phi_t(y'_{[m]}(t,\mu),\mu) \geq \Phi_{t'}(y'_{[m]}(t',\mu),\mu)$$

when $0 < t' < t$, and the limit $\lim_{t\downarrow 0} \Phi_t(y'_{[m]}(t,\mu),\mu)$ exists and its value is just y'^*_{m+1}, namely the optimal value of (ALP)(μ) or (ALD)(μ). It may be verified that

$$\lim_{t\downarrow 0} \mathrm{d_H}(X'_\mu(t), X'_\mu) = 0, \quad \lim_{t\downarrow 0} \mathrm{d_H}(Y'_\mu(t), Y'_\mu) = 0, \tag{23}$$

where $\mathrm{d_H}(\cdot,\cdot)$ denotes the Hausdorff distance. For instance, we demonstrate the first equality as follows. It is obvious that

$$X'_\mu \subset \bigcap_{t>0} X'_\mu(t) (\equiv \hat{X}_\mu),$$

and we only need to prove the opposite inclusion. Let $x' \in \hat{X}_\mu$ then $x' \in \mathcal{F}'_P(\mu)$ and

$$\bar{c}(\mu)^T x' \leq -\Phi_t(y'_{[m]}(t,\mu),\mu) + t\log(n+1), \forall\, t > 0.$$

Taking $t \downarrow 0$ in the above inequality, we obtain

$$\bar{c}(\mu)^T x' \leq y'^*_{m+1},$$

which means that $x' \in X'_\mu$ from the duality theory. From Aubin and Frankowska (1990), we have $\lim_{t\downarrow 0} X'_\mu(t) = \hat{X}_\mu$ and $\lim_{t\downarrow 0} \mathrm{d_H}(X'_\mu(t), X'_\mu) = 0$. We can prove $\lim_{t\downarrow 0} \mathrm{d_H}(Y'_\mu(t), Y'_\mu) = 0$ in the same way.

It follows from (16) and (17) that $(x'(t,\mu), y'(t,\mu)) \in X'_\mu(t) \times Y'_\mu(t)$. Noticing that

$$\mathrm{dist}(x'(t,\mu), X'_\mu) \leq \mathrm{dist}(x'(t,\mu), X'_\mu(t)) + \mathrm{d_H}(X'_\mu(t), X'_\mu) = \mathrm{d_H}(X'_\mu(t), X'_\mu)$$

and

$$\mathrm{dist}(y'(t,\mu), Y'_\mu) \leq \mathrm{dist}(Y'(t,\mu), Y'_\mu(t)) + \mathrm{d_H}(Y'_\mu(t), Y'_\mu) = \mathrm{d_H}(Y'_\mu(t), Y'_\mu),$$

we obtain (16) from (17) directly. The proof is completed. □

From Theorem 2, if $(\tilde{x}'_\mu, \tilde{y}'_\mu)$ is a cluster point of $(x'(t,\mu), y'(t,\mu))$ defined by Theorem 2, $\{t_k\} \subset \mathbf{R}_+$ satisfies $t_k \downarrow 0$ and

$$\tilde{x}'_\mu = \lim_{k\to\infty} x'(t_k,\mu),\ \tilde{y}'_\mu = \lim_{k\to\infty} y'(t_k,\mu),$$

then $(\tilde{x}'_\mu, \tilde{y}'_\mu) \in X'_\mu \times Y'_\mu$. Let

$$\tilde{s}'_\mu \equiv \begin{bmatrix} c \\ -\mu \end{bmatrix} - \begin{bmatrix} A^T \\ -b^T \end{bmatrix} (\tilde{y}'_\mu)_{[m]} - (\tilde{y}'_\mu)_{m+1} e_{[n+1]}.$$

Then

$$\tilde{s}'_\mu \geq 0,\ (\tilde{s}'_\mu)_i(\tilde{x}'_\mu)_i = 0 \ \text{ for } \ i = 1, \ldots, n+1.$$

Since

$$\begin{aligned} x'_i(t_k,\mu) &= \exp(-t_k^{-1} s'_i(t_k,\mu)), \\ s'_i(t_k,\mu) &= c_i - A_i^T y'_{[m]}(t_k,\mu) + \Phi_{t_k}(y'_{[m]}(t_k,\mu),\mu), \end{aligned}$$

where $c_{n+1} \equiv -\mu$ and $A_{n+1} \equiv -b$, we have that if $(\tilde{x}'_\mu)_i > 0$, then

$$s'_i(t_k,\mu) = \mathrm{O}(t_k) \quad (\text{as} \quad t_k \downarrow 0)$$

and if $(\tilde{x}'_\mu)_i = 0$ and $(\tilde{s}'_\mu)_i \geq 0$, then

$$\frac{t_k}{s'_i(t_k,\mu)} \longrightarrow 0 \quad (\text{as} \quad t_k \downarrow 0).$$

Corollary 1. The set $\mathcal{C}_t$ defined by (2) is a solution trajectory to (ALP)(μ) and (ALD)(μ) in the sense

$$\mathrm{dist}[\mathcal{C}_t, \mathcal{F}^*_\mu] \longrightarrow 0 \quad \text{as } t \downarrow 0.$$

3. An unconstrained approach for solving (ALP)(μ)

In this section, we propose an unconstrained approach for solving (ALP)(μ), which is a damped Newton method for minimizing $\Phi_t(y'_{[m]},\mu)$. From Section 2, we know that $\xi'(t,\mu) \equiv \xi'(y'_{[m]}(t,\mu),t)$ satisfies that

$$(\xi'(t,\mu), (y'_{[m]}(t,\mu), -\Phi_t(y'_{[m]}(t,\mu),\mu)), -t\log\xi'(t,\mu)) \in \mathcal{C}_t, t > 0,$$

where $y'_{[m]}(t,\mu)$ is the unique solution to

$$\min \quad \Phi_t(y'_{[m]},\mu). \tag{24}$$

Our algorithm is based on solving (24), which could be used to obtain a point on $\mathcal{C}_t$. Hence if t is small, the algorithm actually finds an approximation solution to ALP(μ) and ALD(μ).

Let

$$\begin{aligned} W(y'_{[m]},t) &= \tilde{A}E(y'_{[m]},t)\tilde{A}^T, \\ p(y'_{[m]},t) &= W(y'_{[m]},t)^{-1}\tilde{A}\xi'(y'_{[m]},t), \\ \eta(y'_{[m]},t) &= (\tilde{A}\xi'(y'_{[m]},t))^T p(y'_{[m]},t). \end{aligned}$$

Then

$$\nabla^2_{y'_{[m]}}\Phi_t(y'_{[m]},t) = \frac{1}{t}\left[W(y'_{[m]},t) - \tilde{A}\xi'(y'_{[m]},t)(\tilde{A}\xi'(y'_{[m]},t))^T\right]$$

and

$$\left[\nabla^2_{y'_{[m]}} \Phi_t(y'_{[m]}, t)\right]^{-1} = t \left[W(y'_{[m]}, t)^{-1} + \frac{p(y'_{[m]}, t) p(y'_{[m]}, t)^T}{1 - \eta(y'_{[m]}, t)} \right].$$

The Newton direction $d^N(y'_{[m]}, t)$ of $\Phi_t(y'_{m]}, \mu)$ at $y'_{[m]}$ is

$$\begin{aligned} d^N(y'_{[m]}, t) &= -\left[\nabla^2_{y'_{[m]}} \Phi_t(y'_{[m]}, t)\right]^{-1} \nabla_{y'_{[m]}} \Phi_t(y'_{[m]}, t) \\ &= -\left[\nabla^2_{y'_{[m]}} \Phi_t(y'_{[m]}, t)\right]^{-1} \tilde{A}\xi'(y'_{[m]}, t) \\ &= -t \left[1 + \eta(y'_{[m]}, t) / \left(1 - \eta(y'_{[m]}, t)\right)\right] p(y'_{[m]}, t) \\ &= \left[-t / \left(1 - \eta(y'_{[m]}, t)\right)\right] p(y'_{[m]}, t), \end{aligned}$$

which is a vector parallel to $-p(y'_{[m]}, t)$. For simplicity, we just take $-p(y'_{[m]}, t)$ as the search direction at $y'_{[m]}$, i.e.,

$$d(y'_{[m]}, t) = -p(y'_{[m]}, t). \tag{25}$$

Obviously, we have

$$\begin{aligned} &\nabla_{y'_{[m]}} \Phi_t(y'_{[m]}, \mu)^T d(y'_{[m]}, t) = -\eta(y'_{[m]}, t), \\ &\nabla_{y'_{[m]}} \Phi_t(y'_{[m]}, \mu)^T d^N(y'_{[m]}, t) = -t\eta(y'_{[m]}, t)/(1 - \eta(y'_{[m]}, t)). \end{aligned}$$

Lemma 2. The vector $y'_{[m]}$ is the minimizer of $\Phi_t(y'_{[m]}, \mu)$ if and only if $\eta(y'_{[m]}, t) = 0$.

Proof. $y'_{[m]}$ is the minimizer of $\Phi_t(y'_{[m]}, \mu)$ if and only if $\nabla_{y'_{[m]}} \Phi_t(y'_{[m]}, t) = \tilde{A}\xi'(y'_{[m]}, t) = 0$, which is equivalent to $\eta(y'_{[m]}, t) = 0$ because of the positive definiteness of $E(y'_{[m]}, t)$. □

Now we are ready to state our algorithm.

Algorithm 1

Step 1 Select an initial point $y'^0_{[m]} \in \mathbf{R}^m$ and parameters $\rho \in (0, 1/2)$ and $\beta \in (0, 1)$. Set $k := 0$.

Step 2 If $\|\tilde{A}\xi'(y'^k_{[m]}, t)\| = 0$, then stop. Otherwise, go to Step 3.

Step 3 Compute

$$d^k = -p(y'^k_{[m]}, t).$$

Step 4 Let l be the smallest nonnegative integer such that

$$\Phi_t(y'_{[m]} + \beta^l d^k, \mu) - \Phi_t(y'^k_{[m]}, \mu) \leq -\beta^l \rho \eta(y'^k_{[m]}, t)$$

and set $y'^{k+1}_{[m]} := y'^k_{[m]} + \alpha_k d^k$, where $\alpha_k := \beta^l$. Return to Step 2 with k replaced by $k+1$.

Theorem 3. The sequence $\{y'^k_{[m]}\}$ generated by Algorithm 1 converges to the solution to (24).

Proof. Since $\{y'^k_{[m]}\} \subset L(y'^0_{[m]})$, we have from Remark 1 that $\{y'^k_{[m]}\}$ is bounded. Let $y'^*_{[m]}$ be any accumulation point of $\{y'^k_{[m]}\}$, we prove that $y'^*_{[m]}$ is just the unique solution to (24). Suppose that this is not true, then

$$\begin{array}{l} \nabla_{y'_{[m]}} \Phi_t(y'^*_{[m]}, \mu) \neq 0, d^* \equiv -p(y'^*_{[m]}, t) \neq 0, \\ \nabla_{y'_{[m]}} \Phi_t(y'^*_{[m]}, \mu)^T d^* = -\eta(y'^*_{[m]}, t) < 0. \end{array} \tag{26}$$

Since

$$\Phi_t(y'^k_{[m]} + \alpha_k d^k) - \Phi_t(y'^k_{[m]}) \leq -\alpha_k \rho \eta(y'^k_{[m]}, t),$$

we obtain that $\{\alpha_k \eta(y'^k_{[m]}, t)\}$ converges to 0. In order to prove $\eta(y'^k_{[m]}, t) \longrightarrow 0$, we show that $\{\alpha_k\}$ is bounded away from 0. Now suppose that there exists a subsequence of $\{\alpha_k\}$, say $\{\alpha_{k_i}\}$, tending to 0. By the line search rule, we have

$$\frac{\Phi_t(y'^{k_i}_{[m]} + \beta^{-1}\alpha_{k_i} d^{k_i}) - \Phi_t(y'^{k_i}_{[m]})}{\beta^{-1}\alpha_{k_i}} > -\rho \eta(y'^{k_i}_{[m]}, t). \tag{27}$$

Since $\alpha_{k_i} \to 0$, taking the limit of both sides of (26) yields

$$\nabla_{y'_{[m]}} \Phi_t(y'^*_{[m]}, \mu)^T d^* \geq -\rho \eta(y'^*_{[m]}, t),$$

i.e.,

$$-(1-\rho)\eta(y'^*_{[m]}, t) \geq 0.$$

Since $\rho \in (0, 1/2)$, the above inequality implies $-\eta(y'^*_{[m]}, t) \geq 0$, which contradicts with (26). Thus $\{\alpha_k\}$ is bounded away from 0 and $\{\eta(y'^k_{[m]}, t)\}$ converges to 0. That is

$$\lim_{k\to\infty} -\nabla_{y'_{[m]}} \Phi_t(y'^k_{[m]}, \mu)^T d^k = -\nabla_{y'_{[m]}} \Phi_t(y'^*_{[m]}, \mu)^T d^* = \eta(y'^*_{[m]}, t) = 0.$$

Thus, we have from Lemma 2 that $y'^*_{[m]}$ is the unique solution to (24). Since every accumulation point of $\{y'^k_{[m]}\}$ is the unique solution to (24) and $\{y'^k_{[m]}\}$ is bounded, we have that $\{y'^k_{[m]}\}$ converges to the solution to (24). The proof is completed. □

Algorithm 1 is a simplified version of the Newton method for solving (24). In order to prove superlinear convergence rate, we slightly sharpen the algorithm.

A Revised Version of Algorithm 1 — Algorithm 2

Step 1 Select an initial point $y'^0_{[m]} \in \mathbf{R}^m$ and parameters $\rho \in (0, 1/2)$ and $\beta \in (0, 1)$. Set $k := 0$.

Step 2 If $\|\tilde{A}\xi'(y'^k_{[m]}, t)\| = 0$, then stop. Otherwise, go to Step 3.

Step 3 Compute

$$d^k_N = -tp(y'^k_{[m]}, t)/(1 - \eta(y'^k_{[m]}, t)).$$

Step 4 Let l be the smallest nonnegative integer such that

$$\Phi_t(y'_{[m]} + \beta^l d^k_N, \mu) - \Phi_t(y'^k_{[m]}, \mu) \leq -\beta^l \rho t\eta(y'^k_{[m]}, t)/(1 - \eta(y'^k_{[m]}, t))$$

and set $y'^{k+1}_{[m]} := y'^k_{[m]} + \alpha_k d^k$, where $\alpha_k := \beta^l$. Return to Step 2 with k replaced by $k + 1$.

Corollary 2. The sequence $\{y'^k_{[m]}\}$ generated by Algorithm 2 converges to the solution to (24).

Proof. Similar to the proof of Theorem 3. □

Lemma 3. For all k sufficiently large, the stepsize $\alpha_k = 1$ is chosen in Step 4 of Algorithm 2.

Proof. Since $\{y'^k_{[m]}\}$ converges to the solution to (24), one has that for k sufficiently large,

$$\begin{aligned}\Phi_t(y'^k_{[m]} + \alpha_k d^k_N, \mu) = \Phi_t(y'^k_{[m]}, \mu) + \alpha_k \nabla_{y'_{[m]}} \Phi_t(y'^k_{[m]}, \mu)^T d^k_N + \\ +(\alpha_k^2/2) d^{kT}_N \nabla^2_{y'_{[m]}} \Phi_t(y'^k_{[m]}, \mu)^T d^k_N + \mathrm{o}(\|\alpha_k d^k_N\|^2).\end{aligned}$$

It follows from the definition of d^k_N that

$$\begin{aligned}\Phi_t(y'^k_{[m]} + \alpha_k d^k_N, \mu) &= \Phi_t(y'^k_{[m]}, \mu) - t\alpha_k \eta(y'^k_{[m]}, t)/(1 - \eta(y'^k_{[m]}, t)) + \\ &\quad + t\eta(y'^k_{[m]}, t)\alpha_k^2/2(1 - \eta(y'_{[m]}, t)) + \mathrm{o}(\|\alpha_k d^k_N\|^2) \\ &= \Phi_t(y'^k_{[m]}, \mu) - (\alpha_k - \alpha_k^2/2)t\eta(y'^k_{[m]}, t)/(1 - \eta(y'^k_{[m]}, t)) + \\ &\quad + \mathrm{o}(\|\alpha_k d^k_N\|^2)\end{aligned}$$

Since $\{y'^k_{[m]}\} \longrightarrow y'^*_{[m]}$, where $y'^*_{[m]}$ is the solution to (24), there exists $\tau > 0$ such that

$$\lambda_{\min}[E(y'^*_{[m]}, t)] \geq \tau \tag{28}$$

for k sufficiently large, where $\lambda_{\min}$ stands for the smallest eigenvalue. Noting the definition of d^k_N and (28), we obtain

$$\begin{aligned}
\|d^k_N\|^2 &= (t^2\eta(y'^k_{[m]}, t)^2/(1-\eta(y'^k_{[m]}, t))^2)\|p(y'^k_{[m]}, t)\|^2 \\
&= t^2\eta(y'^k_{[m]}, t)^2\nabla\Phi_t(y'^k_{[m]}, \mu)^T E(y'^k_{[m]}, t)^{-2}\nabla\Phi_t(y'^k_{[m]}, \mu)/(1-\eta(y'^k_{[m]}, t))^2 \\
&= t^2\eta(y'^k_{[m]}, t)^2\nabla\Phi_t(y'^k_{[m]}, \mu)^T E(y'^k_{[m]}, t)^{-1/2} E(y'^k_{[m]}, t)^{-1} \times \\
&\qquad \times E(y'^k_{[m]}, t)^{-1/2}\nabla\Phi_t(y'^k_{[m]}, \mu)/(1-\eta(y'^k_{[m]}, t))^2 \\
&\leq \tau^{-1}t^2\eta(y'^k_{[m]}, t)^3/(1-\eta(y'^k_{[m]}, t))^2.
\end{aligned}$$

For k sufficiently large, we have

$$\frac{1}{2} - \frac{\tau t\eta(y'^k_{[m]}, t)^2}{1-\eta(y'^k_{[m]}, t)} \geq \rho. \tag{29}$$

It follows from

$$\begin{aligned}
&\Phi_t(y'^k_{[m]} + \alpha_k d^k_N, \mu) - \Phi_t(y'^k_{[m]}, \mu) \leq \\
\leq \quad & -\left[\alpha_k - \alpha_k^2/2 - \alpha_k^2\tau t\eta(y'^k_{[m]}, t)^2/(1-\eta(y'^k_{[m]}, t))\right]\left[t\eta(y'^k_{[m]}, t)/(1-\eta(y'^k_{[m]}, t))\right]
\end{aligned}$$

that α_k must be 1 because $\rho \in (0, 1/2)$ and

$$1 - \frac{1}{2} - \frac{\tau t\eta(y'^k_{[m]}, t)^2}{1-\eta(y'^k_{[m]}, t)} \geq \rho$$

holds for k sufficiently large. The proof is completed. □

Theorem 4. The sequence $\{y'^k_{[m]}\}$ generated by Algorithm 2 converges to the solution to (24) at a superlinear rate.

Proof. From Lemma 3, we know that $\alpha_k \equiv 1$ for k sufficiently large and $y'^{k+1}_{[m]} = y'^k_{[m]} + d^k_N$. Let $y'^*_{[m]}$ be the solution to (24), then

$$\begin{aligned}
\|y'^{k+1}_{[m]} - y'^*_{[m]}\| &= \|y'^k_{[m]} - (\nabla^2\Phi_t(y'^k_{[m]}, \mu))^{-1}\nabla\Phi_t(y'^k_{[m]}, \mu) - y'^*_{[m]}\| \\
&= \|(\nabla^2\Phi_t(y'^k_{[m]}, \mu))^{-1}[\nabla\Phi_t(y'^*_{[m]}, \mu) - \\
&\qquad -\nabla\Phi_t(y'^k_{[m]}, \mu) + \nabla^2\Phi_t(y'^k_{[m]}, \mu)(y'^k_{[m]} - y'^*_{[m]})]\| \\
&= \left\|(\nabla^2\Phi_t(y'^k_{[m]}, \mu))^{-1}(\int_0^1(\nabla^2\Phi_t(y'^k_{[m]} + s(y'^*_{[m]} - y'^k_{[m]}), \mu) - \right. \\
&\qquad \left. -\nabla^2\Phi_t(y'^k_{[m]}, \mu))(y'^*_{[m]} - y'^k_{[m]})\mathrm{d}s\right\|.
\end{aligned}$$

Since $y'^{k}_{[m]} \longrightarrow y'^{*}_{[m]}$, one has, for k sufficiently large, $\alpha_k \equiv 1$ and

$$\|(\nabla^2\Phi_t(y'^{k}_{[m]}, \mu))^{-1}\| \leq c,$$

for some positive constant c. Therefore,

$$\left\| y'^{k+1}_{[m]} - y'^{*}_{[m]} \right\| \leq$$
$$c \left\| \int_0^1 \left[\nabla^2\Phi_t \left(y'^{k}_{[m]} + s(y'^{*}_{[m]} - y'^{k}_{[m]}), \mu \right) - \nabla^2\Phi_t(y'^{k}_{[m]}, \mu) \right] \mathrm{d}s \right\| \left\| y'^{*}_{[m]} - y'^{k}_{[m]} \right\|.$$

The superlinear convergence of $\{y'^{k}_{[m]}\}$ comes from the above inequality and the continuity of $\nabla^2\Phi_t(\cdot, \mu)$. The proof is completed. □

4. Concluding remarks

We define a new central path by a different perturbation on the right hand side term of the complementarity equation in the KKT system of the Karmarkar form of LP. Similar to the traditional central path, the new central path defines a trajectory toward the solution of the Karmarkar form of LP as the parameter approaches zero. A damped Newton method is shown to converge to any given point on this central path at a superlinear rate; therefore provides a method for solving the LP problem if the parameter is set small enough to a user-specified tolerance.

It is interesting to note that the points on the traditional central path are the solutions to the log-barrier problem of the standard LP while the points on the new central path can be interpreted as the solution to the log-exp problem of the Karmarkar LP. By introducing self-concordance, Nesterov and Nemirovskii (1994) show that Newton's method can be used to efficiently "follow" the central path, hence producing polynomial algorithms. This paper did not touch the topic of how to follow the new central path but it is certainly a reasonable future step of research. The proposed algorithm, on the other hand, is globally convergent to a point on the central path with arbitrarily small t, therefore it is more like a barrier method with a pre-specified parameter. Another possible direction of research is to understand the computational impact of such a central path. Since the log-exp function is a smooth approximation of the vecmax function, it looks that the proposed method tends to reduce the largest slack in the dual problem. Therefore the proposed method might be quite robust for badly-scaled problems. However, no conclusion can be made unless enough computational evidence is provided.

References

Aubin, J. P. and H. Frankowska. Set-Valued Analysis. Birkhäuser, Boston, 1990.

Chen, C. and O. L. Mangasarian. A class of smoothing functions for nonlinear and mixed complementarity problems. *Comput. Optim. Appli.*, 5:97–138, 1996.

Fang S. C. A new unconstrained convex programming approach to linear programming. *Zeitschrift für Operations Research*, 36:149–161, 1992.

Li X. S. A aggregate constraint method for nonlinear programming. *J. Oper. Resear. Soc.*, 42: 1003-1010, 1991.

Li X. S. An entropy-based aggregate method for minimax optimization. *Eng. Optim.*, 18:227–285, 1992.

Nesterov, Y. and A. Nemirovskii. Interior Point Polynomial Algorithms in Convex Programming. SIAM Publications, Philadelphia, PA, USA,1994.

Peng, J. M. and Z. H. Lin. A non-interior continuation method for generalized linear complementarity problems. *Math. Program., Ser. A*, 86:533-563, 1999.

Rockafellar, R. T. and R. J. B. Wets. Variational Analysis. Springer-Verlag Berlin Heidelberg, 1998.

Templeman A. B. and X. S. Li. A maximum approach to constrained nonlinear programming. *Eng. Optim.*, 12: 191–205, 1987.

Ye, Y. Interior Point Algorithms: Theory and Analysis. John Wiley & Sons, Inc., 1997.