Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 14, No. 2, pp. 261–272, 2018

ANALYSIS OF THE PRIMAL AFFINE SCALING CONTINUOUS TRAJECTORY FOR CONVEX PROGRAMMING*

Xun Qian and Li-Zhi Liao

Abstract: A weighted first-order primal affine scaling continuous trajectory for linearly constrained smooth convex programming is studied in this paper. By assuming the existence of an optimal solution in the linear case or the boundness of the optimal solution set in the general case, we show that starting from any interior feasible point, (i) every accumulation point is indeed an optimal solution; and (ii) if the objective function is analytic, the primal affine scaling continuous trajectory converges to a point which is actually in the relative interior of the optimal solution set. As we know, this result is the first one to obtain the convergence of the primal affine scaling continuous trajectory in the nonlinear case for linearly constrained convex programming.

Key words: *affine scaling, continuous trajectory, interior point method, convex programming, ordinary differential equation*

Mathematics Subject Classification: *90C25, 90C51, 37N40*

1 Introduction

Consider the following linearly constrained convex programming problem

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & Ax = b,\ x \geq 0, \end{array} \tag{P}$$

where $f(x)$ is a smooth convex function, $b \in R^m$, and A is an $m \times n$ matrix with full row rank, $m < n$. As a blanket assumption, we assume that the optimal value for problem (P) is finite and attainable, therefore, we use min rather than inf in problem (P).

The following notations are used in the sequel

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

where $\mathcal{P}^{++}$ is called the relative interior of $\mathcal{P}^+$. It is conventional to assume that $\mathcal{P}^{++}$ is nonempty in the analysis of interior point methods. We also assume $f(x) \in C^2$ on R^n.

The affine scaling method was first proposed by Dikin [6] in 1967, then many researchers have made numerous contributions since, for instance, the affine scaling algorithms by Dikin

*This research was supported in part by grants from Hong Kong Baptist University (FRG) and the General Research Fund (GRF) of Hong Kong.

[7], Saigal [22], Tseng and Luo [27], Tsuchiya [28, 29], Monteiro and Tsuchiya [18], Gonzaga and Carlos [11], Sun [24, 25], Tseng *et al.* [26], and so on. The affine scaling continuous trajectory was also studied for linear programming, for example, by Adler and Monteiro [1], Bayer and Lagarias [3], Liao [12], Megiddo and Shub [16], Monteiro [17], and so on. It should be noted that for linear programming, the primal affine scaling continuous trajectory actually contains the primal central path if the initial point is on the central path. In the interior point method, the central path plays a vital role. For the central path where the barrier function $\sum_{i=1}^{n} -\ln x_i$ is used in problem (P), convergence can be obtained under the strictly complementarity condition [15], the analyticity of $f(x)$ [19], or the condition that there exists a subspace W of R^n such that $\mathrm{Ker}(\nabla^2 f(x)) = W$ [8]. Some ill-behaved central path examples in convex optimization can be found in [10], however these examples are different from problem (P).

In this paper, we are interested in the first-order primal affine scaling continuous trajectory, which has already been studied for linear programming in [1], but not yet for convex programming (P). Compared with [1], in the linear case, we do not require the boundness of the optimal solution set, instead, we only need the existence of a finite optimal solution. It should be noted that the first-order primal affine scaling trajectory is contained in the Cauchy trajectories for convex semidefinite programming [13], but there is no strong convergence result for the Cauchy trajectories there. To our knowledge, our result here is the first one to obtain the strong convergence of the primal affine scaling continuous trajectory in the nonlinear case for problem (P).

For simplicity, in what follows, $\|\cdot\|$ denotes the 2-norm. C^k stands for the class of kth order continuously differentiable functions. Unless otherwise specified, x_j denotes the jth component of a vector x, e denotes the column vector of all ones, and e_i denotes the unity column vector whose ith component is 1, the dimensions of e and e_i are clear from the context. For any index subset $J \subseteq \{1, \ldots, n\}$, we denote by x_J the vector composed of those components of $x \in R^n$ indexed by $j \in J$, $\mathrm{rank}\,(Q)$ denotes the rank of the matrix Q.

The rest of this paper is organized as follows. In Section 2, we introduce the corresponding ordinary differential equation (ODE) system for the weighted first-order primal affine scaling continuous trajectory, verify that the ODE system has a unique solution in $[t_0, +\infty)$, and show some properties of the primal affine scaling continuous trajectory. In Section 3, we prove that every accumulation point of the primal affine scaling continuous trajectory is an optimal solution for problem (P). Finally, in Section 4, we show the strong convergence of the weighted first-order primal affine scaling continuous trajectory under the condition that $f(x)$ is analytic.

2 Properties of the Continuous Trajectory

The weighted first-order primal affine scaling direction for problem (P) can be given by

$$-DP_{AD}D\nabla f(x),$$

where $x \in R^n_{++}$, $X = \mathrm{diag}\,(x) \in R^{n\times n}$, $w \in R^n_{++}$ (a given vector), $W = \mathrm{diag}\,(w) \in R^{n\times n}$, $D = W^{-\frac{1}{2}}X$, $P_{AD} = I_n - DA^T(AD^2A^T)^{-1}AD$, and I_n (or I) stands for the $n \times n$ identity matrix. The first-order primal affine scaling direction usually adopts $w \equiv e$. As a result, the weighted first-order primal affine scaling continuous trajectory for problem (P) is the solution curve of the following ODE system

$$\frac{dx}{dt} = -DP_{AD}D\nabla f(x),\ x(t_0) = x^0 \in \mathcal{P}^{++},\ t \geq t_0 > 0. \tag{2.1}$$

The following assumptions are made throughout this paper.

Assumption 2.1. If $f(x) = c^T x$, we assume that there exists a point $x^* \in \mathcal{P}^+$ such that $c^T x^*$ is the optimal value of problem (P). Otherwise, we assume that the optimal solution set of problem (P) is non-empty and bounded.

Assumption 2.2. The set $\mathcal{P}^{++}$ is not empty.

First we state two simple technical lemmas without proof.

Lemma 2.1. $(AD^2A^T)^{-1} \in C^1$ *on* R^n_{++}.

Lemma 2.2. $DP_{AD}D\nabla f(x) \in C^1$ *on* R^n_{++}.

Lemma 2.2 reveals the smoothness property for the right-hand side of ODE system (2.1). Theorem 2.3 and Theorem 2.4 below guarantee the existence, uniqueness, and feasibility for the solution of ODE system (2.1).

Theorem 2.3. *For ODE system (2.1), there exists a unique solution $x(t)$ with a maximal existence interval $[t_0, \alpha)$, in addition, $x(t) > 0$ on this existence interval.*

Proof. By Lemma 2.2, $DP_{AD}D\nabla f(x)$ is locally Lipschitz continuous on R^n_{++}. Since R^n_{++} is an open set, from the Cauchy-Peano theorem and Picard-Lindelöf theorem, there exists a unique solution $x(t)$ of ODE system (2.1) on the maximal existence interval $[t_0, \alpha)$, for some $\alpha > t_0$ or $\alpha = +\infty$.

Because the right-hand side of ODE system (2.1) is defined on the open set $(0, +\infty) \times R^n_{++}$, the solution of ODE system (2.1) is of course in the open set R^n_{++}, so $x(t)$ is positive on the existence interval. The proof is complete. □

Later in this section, it will be shown that $\alpha = +\infty$ (Theorem 2.6). To simplify the following presentation, in the remaining of this paper, $x(t)$ (or $X(t)$) will be replaced by x (or X) whenever no confusion would occur.

Theorem 2.4. *Let $x(t)$ be the solution of ODE system (2.1) with the maximal existence interval $[t_0, \alpha)$. Then $Ax(t) = b$ $\forall t \in [t_0, \alpha)$.*

Proof. We know that for any $t \in [t_0, \alpha)$

$$x(t) = x^0 - \int_{t_0}^{t} (DP_{AD}D\nabla f(x)|_{t=\tau})d\tau.$$

Noticing

$$ADP_{AD} = AD - AD^2A^T(AD^2A^T)^{-1}AD \equiv 0,$$

we can get

$$Ax(t) = Ax^0 - \int_{t_0}^{t} (ADP_{AD}D\nabla f(x)|_{t=\tau})d\tau = b.$$

Thus the theorem is proved. □

Next we show that the solution curve is contained in a bounded set.

Theorem 2.5. *The unique solution $x(t)$ of ODE system (2.1) is contained in a bounded set in R^n_+.*

Proof. If $f(x) = c^T x$, from Theorem 2.3, $x(T) > 0$ for any $T \in [t_0, \alpha)$, then we can define

$$V_1(t) = \sum_{i=1}^{n} \frac{w_i(x(T)_i - x_i^*)}{x(t)_i}, \quad t \in [t_0, \alpha),$$

where x^* is from Assumption 2.1. From Theorem 2.4, we have

$$\begin{aligned} \frac{dV_1(t)}{dt} &= \sum_{i=1}^{n} -\frac{w_i(x(T)_i - x_i^*)}{x(t)_i^2} \cdot \frac{dx_i}{dt} \\ &= -(x(T) - x^*)^T W X(t)^{-2} \frac{dx}{dt} \\ &= (x(T) - x^*)^T D^{-2} D P_{AD} D c \\ &= (x(T) - x^*)^T D^{-2} (D^2 - D^2 A^T (AD^2 A^T)^{-1} AD^2) c \\ &= (x(T) - x^*)^T c - (b - b)^T (AD^2 A^T)^{-1} AD^2 c \\ &= c^T x(T) - c^T x^* \\ &\geq 0, \end{aligned}$$

then

$$V_1(t_0) \leq V_1(T) = \sum_{i=1}^{n} w_i - \sum_{i=1}^{n} \frac{w_i x_i^*}{x(T)_i} \leq \sum_{i=1}^{n} w_i. \tag{2.2}$$

Therefore for any $T \in [t_0, \alpha)$, we have

$$\|x(T)\| \leq e^T x(T) \leq \frac{\max_i(x_i^0)}{\min_i(w_i)} \sum_{i=1}^{n} \frac{w_i x(T)_i}{x_i^0} \leq \frac{\max_i(x_i^0)}{\min_i(w_i)} \left(\sum_{i=1}^{n} w_i + \sum_{i=1}^{n} \frac{w_i x_i^*}{x_i^0} \right).$$

The last inequality is from (2.2), which indicates that $x(T)$ is bounded, and the bound depends only on x^0, x^*, and w.

Otherwise, noticing for $t \in [t_0, \alpha)$,

$$\frac{df(x(t))}{dt} = -\nabla f(x)^T D P_{AD} D \nabla f(x) = -\|P_{AD} D \nabla f(x)\|^2 \leq 0,$$

we know that $f(x(t))$ is a nonincreasing function on $t \in [t_0, \alpha)$. Hence $x(t)$ will be contained in the level set $\{x \mid x \in \mathcal{P}^+,\ f(x) \leq f(x^0)\}$. Under Assumption 2.1, from Theorem 24 on page 93 in [9], the level set will be bounded as well. Thus the proof is complete. □

After we get the boundedness of the solution curve, we can extend the existence interval of the solution to infinity.

Theorem 2.6. *Let $x(t)$ be the solution of ODE system (2.1) with the maximal existence interval $[t_0, \alpha)$. Then $\alpha = +\infty$.*

Proof. Assume $\alpha \neq +\infty$. From Theorem 2.5, we know that there exists an $M > 0$ such that $0 < x(t) \leq Me$ $\forall t \in [t_0, \alpha)$. Furthermore, P_{AD} is symmetric and idempotent, which leads to $\|P_{AD}\| \leq 1$. Therefore the vector $P_{AD} D \nabla f(x)$ is bounded. Then we know that there exists an $L > 0$ such that for every $i \in \{1, \ldots, n\}$, we have

$$\left| \frac{dx_i}{dt} \right| \leq L x_i \quad \forall t \in [t_0, \alpha), \tag{2.3}$$

and this L depends only on M, and $f(x)$.

For every $i \in \{1, \ldots, n\}$, from inequality (2.3) and $0 < x(t) \le Me$ $\forall t \in [t_0, \alpha)$, we know that

$$\left| \frac{dx_i}{dt} \right| \le LM \quad \forall t \in [t_0, \alpha), \tag{2.4}$$

furthermore, $x(t)$ is continuous on $[t_0, \alpha)$, and it is not hard to see that $\lim_{t \to \alpha^-} x(t)$ exists. We denote this limit as $x(\alpha)$. Evidently $x(\alpha) \ge 0$. According to the Extension Theorem in §2.5, [2], we know that the solution $x(t)$ will go to the boundary of the open set $(0, +\infty) \times R^n_{++}$. But because of the hypothesis, $\alpha \ne +\infty$, so there must exist at least one $i \in \{1, \ldots, n\}$ such that $x_i(\alpha) = 0$. From inequality (2.3), we know that if $t \in [t_0, \alpha)$,

$$\frac{dx_i}{x_i} \ge -Ldt.$$

Integrating the inequality above, we have for every $t \in [t_0, \alpha)$

$$\ln x_i(t) - \ln x_i(t_0) \ge -L(t - t_0).$$

Since $x_i(t) \to x_i(\alpha) = 0$ as $t \to \alpha^-$, $\ln x_i(t) - \ln x_i(t_0) \to -\infty$ as $t \to \alpha^-$, but $-L(t - t_0) \ge -L(\alpha - t_0)$. This is a contradiction. Therefore $\alpha = +\infty$, and the proof is complete. □

From Theorem 2.6, we can define the limit set for the solution of ODE system (2.1). Let $x(t)$ be the solution of ODE system (2.1), the limit set of $\{x(t)\}$ can be defined as follows

$$\Omega^1(x^0) = \left\{ x \in R^n \mid \exists \ \{t_k\}_{k=0}^{+\infty} \text{ with } \lim_{k \to +\infty} t_k = +\infty \ \text{ such that } \lim_{k \to +\infty} x(t_k) = x \right\}.$$

Theorem 2.7. *The limit set $\Omega^1(x^0)$ is nonempty, compact, and connected. Furthermore $\Omega^1(x^0)$ is contained in $\mathcal{P}^+$.*

Proof. From Theorems 2.3, 2.4, and 2.6, we know that the limit set $\Omega^1(x^0)$ is contained in $\mathcal{P}^+$. From Theorem 2.5, we know that the solution $x(t)$ is contained in a bounded closed set. So similar to the proof of Theorem 1.1 on page 390 in [5] (the proof in [5] is for $n = 2$, but it can be easily exntended to the general case), it can be verified that $\Omega^1(x^0)$ is nonempty, compact, and connected. □

Lemma 2.8. *([4]) Suppose f is differentiable (i.e., its gradient ∇f exists at each point in dom$f$). Then f is convex if and only if dom$f$ is convex and*

$$f(y) \ge f(x) + \nabla f(x)^T (y - x) \tag{2.5}$$

holds for all x, $y \in$ dom$f$.

Now we introduce a kind of potential function for ODE system (2.1). In 1983, Losert and Akin [14] introduced a kind of potential function for both the discrete and continuous dynamical systems in a classical model of population genetics. Their potential function can be extended for our purpose. The potential function $I(x, y)$ for ODE system (2.1) can be defined as

$$I(x, y) = \sum_{i=1}^{n} (w_i \ln x_i) + \sum_{i=1}^{n} w_i \frac{y_i}{x_i}, \tag{2.6}$$

where $x \in R^n_{++}$ is the variable, $y \in R^n_+$ is a parameter.

In the rest of this section, we will show the weak convergence for the solution of ODE system (2.1), i.e., $DP_{AD} D \nabla f(x) \to 0$ as $t \to +\infty$. But first, we reveal some fundamental results for the solution of ODE system (2.1).

Theorem 2.9. *Let $x(t)$ be the solution of ODE system (2.1). Then $f(x(t))$ is a nonincreasing function on $[t_0,+\infty)$. Furthermore, if $x^0 \in P^{++}$ is an optimal solution for problem (P), then $x(t) \equiv x^0$ on $[t_0,+\infty)$; otherwise $f(x(t))$ is a strictly decreasing function on $[t_0,+\infty)$.*

Proof. Since for $t \geq t_0$,

$$\frac{df(x(t))}{dt} = -\|P_{AD}D\nabla f(x)\|^2 \leq 0,$$

we know that $f(x(t))$ is a noincreasing function on $[t_0,+\infty)$.

The KKT conditions for problem (P) can be stated as follows

$$\begin{cases} Ax = b, \ x \geq 0, \\ Xz = 0, \ z \geq 0, \\ A^T y + z = \nabla f(x), \end{cases} \tag{2.7}$$

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

If $x \in \mathcal{P}^{++}$ is an optimal solution, there must exist corresponding (y, z) such that system (2.7) holds, then

$$z = 0 \quad \text{and} \quad A^T y = \nabla f(x),$$

thus, it is easy to see that

$$P_{AD}D\nabla f(x) = P_{AD}DA^T y = 0.$$

So if $x^0 \in \mathcal{P}^{++}$ is an optimal solution, we know that the right-hand side of ODE system (2.1) equals zero at $x = x_0$, i.e.,

$$DP_{AD}D\nabla f(x)|_{x=x_0} = 0,$$

therefore $x(t) \equiv x^0$ for $t \geq t_0$ is a solution of ODE system (2.1). Because of the uniqueness of the solution, we know that $x(t) \equiv x^0$ on $[t_0,+\infty)$.

If $x^0 \in \mathcal{P}^{++}$ is not an optimal solution, we will show $f(x(t))$ is a strictly decreasing function on $[t_0,+\infty)$. If not, then there must exist t_1 and t_2 with $t_0 \leq t_1 < t_2$ such that $f(x(t_1)) = f(x(t_2))$. Since $\frac{df(x(t))}{dt} \leq 0$, we know that when $t_1 \leq t \leq t_2$, $\frac{df(x(t))}{dt} \equiv 0$. From $\frac{df(x(t))}{dt} = -\|P_{AD}D\nabla f(x)\|^2 = -\|D^{-1}\frac{dx}{dt}\|^2 = 0$, Theorem 2.3 and Theorem 2.6, it is easy to see $\frac{dx}{dt} = 0$ on $[t_1, t_2]$, then $x(t) \equiv x(t_1)$ on $[t_0,+\infty)$ will be a solution of the ODE

$$\frac{dx}{dt} = -DP_{AD}D\nabla f(x)$$

that passes through the point $(t_1, x(t_1))$. But according to the uniqueness of the solution, we know that the solution of ODE system (2.1) is actually $x(t) \equiv x(t_1) = x^0$ on $[t_0,+\infty)$. Hence $\frac{dx}{dt}|_{t=t_0} = 0$, which implies

$$(I - A^T(AD_0^2A^T)^{-1}AD_0^2)\nabla f(x^0) = 0,$$

where $D_0 = W^{-\frac{1}{2}}X(t_0)$. Let $y = (AD_0^2A^T)^{-1}AD_0^2\nabla f(x^0)$ and $z = 0$, then (x^0, y, z) will satisfy the KKT system (2.7). Therefore x^0 must be an optimal solution. However this is a contradiction. Thus $f(x(t))$ is a strictly decreasing function on $[t_0,+\infty)$. □

Lemma 2.10. *(Barbalat's Lemma [23]) If the differentiable function $f(t)$ has a finite limit as $t \to +\infty$, and $\dot{f}$ is uniformly continuous, then $\dot{f} \to 0$ as $t \to +\infty$.*

Lemma 2.11. *If $0 < x \leq Me$ with $M > 0$, then for every $i \in \{1, \ldots, n\}$, every entry of*

$$\frac{\partial DP_{AD}D\nabla f(x)}{\partial x_i}$$

is bounded, and the bound depends only on A, M, n, w, and $f(x)$.

Proof. Let $H = (AD^2A^T)^{-1}AD^2$. From Lemma 3 and the Remark in Sun [24], we know that if $x > 0$ every entry of $(AD^2A^T)^{-1}AD^2$ is bounded, and the bound depends only on A and n. Notice

$$\frac{\partial DP_{AD}D}{\partial x_i} = \frac{\partial D^2}{\partial x_i} - \frac{\partial D^2A^T}{\partial x_i}H - D^2A^T\frac{\partial (AD^2A^T)^{-1}}{\partial x_i}AD^2 - H^T\frac{\partial AD^2}{\partial x_i},$$

and

$$D^2A^T\frac{\partial (AD^2A^T)^{-1}}{\partial x_i}AD^2 = -2x_iH^T(AW^{-\frac{1}{2}}e_ie_i^TW^{-\frac{1}{2}}A^T)H.$$

Therefore when $0 < x \leq Me$, every entry of

$$\frac{\partial DP_{AD}D}{\partial x_i}$$

is bounded, and the bound depends only on A, M, w, and n. Then it is evident that for every $i \in \{1, \ldots, n\}$, every entry of

$$\frac{\partial DP_{AD}D\nabla f(x)}{\partial x_i}$$

is bounded, and the bound depends only on A, M, n, w, and $f(x)$. □

We now show the weak convergence for the solution of ODE system (2.1).

Theorem 2.12. *Let $x(t)$ be the solution of ODE system (2.1). Then*

$$\lim_{t\to+\infty} DP_{AD}D\nabla f(x) = 0.$$

Proof. From Theorem 2.5, we know that there exists an $M > 0$ such that the solution $x(t)$ of ODE system (2.1) is contained in the bounded closed set $\{x \in R^n | 0 \leq x \leq Me\}$. This along with Lemma 2.11 indicates that there exists a constant L_1 such that for every $i \in \{1, \ldots, n\}$, every entry of

$$\frac{\partial \nabla f(x)^T DP_{AD}D\nabla f(x)}{\partial x_i} \tag{2.8}$$

is bounded by L_1, and L_1 depends only on A, M, n, w, and $f(x)$.

From Theorem 2.9, we know that $f(x(t))$ is a nonincreasing function and $f(x(t)) \geq f(x^*)$ on $[t_0, +\infty)$. Thus $f(x(t))$ has a finite limit as $t \to +\infty$. From (2.4), we have

$$\begin{aligned}
&\left|\frac{df(x(t))}{dt}\,|_{t=t_1} - \frac{df(x(t))}{dt}\,|_{t=t_2}\right| \\
= &\left|\int_0^1 \frac{\partial \nabla f(x)^T DP_{AD}D\nabla f(x)}{\partial x}|_{x=x(t_2)+\tau(x(t_1)-x(t_2))}(x(t_1)-x(t_2))d\tau\right| \\
\leq &\sqrt{n}L_1 \cdot \|x(t_1) - x(t_2)\|
\end{aligned}$$

$$
\begin{aligned}
&= \sqrt{n}L_1 \cdot \| \int_{t_1}^{t_2} \frac{dx}{d\tau} d\tau \| \\
&\leq nL_1 LM|t_1 - t_2|,
\end{aligned}
$$

where the last inequality is obtained from inequality (2.4).

Thus, $\frac{df(x(t))}{dt}$ is uniformly continuous. From Barbalat's Lemma, we know that

$$\lim_{t\to+\infty} \frac{df(x(t))}{dt} = -\lim_{t\to+\infty} \|P_{AD} D\nabla f(x)\|^2 = 0,$$

which indicates

$$\lim_{t\to+\infty} DP_{AD}D\nabla f(x) = 0.$$

Thus the proof is complete. □

3 Optimality of the Cluster Point(s)

In this section, we show that every accumulation point of the solution of ODE system (2.1) is an optimal solution for problem (P).

Theorem 3.1. *For any $x^{(1)} \in \Omega^1(x^0)$, $x^{(1)}$ is an optimal solution for problem (P).*

Proof. We prove this by contradiction. Assume $x^{(1)}$ is not an optimal solution for problem (P), then from Theorem 2.9, we know $f(x^0) > f(x^{(1)}) = \lim_{k\to+\infty} f(x(t_k)) > f(x^*)$, where $\lim_{k\to+\infty} x(t_k) = x^{(1)}$. Let's define

$$y^{(1)} = \frac{f(x^{(1)}) - f(x^*)}{2(f(x^0) - f(x^*))} x^0 + \left[1 - \frac{f(x^{(1)}) - f(x^*)}{2(f(x^0) - f(x^*))}\right] x^*,$$

then $y^{(1)} \in \mathcal{P}^{++}$. Since $y^{(1)}$ is a convex combination of x^0 and x^*, obviously

$$f(y^{(1)}) \leq \frac{f(x^{(1)}) - f(x^*)}{2(f(x^0) - f(x^*))} f(x^0) + \left[1 - \frac{f(x^{(1)}) - f(x^*)}{2(f(x^0) - f(x^*))}\right] f(x^*) = \frac{f(x^{(1)}) + f(x^*)}{2}.$$

Then we can define

$$V_2(t) = I(x(t), y^{(1)}) = \sum_{i=1}^n w_i (\ln x_i) + \sum_{i=1}^n w_i \frac{y_i^{(1)}}{x_i},$$

where $t \in [t_0, +\infty)$ and $x(t)$ is the unique solution of ODE system (2.1). Then from Theorem 2.4 and Lemma 2.8, we have $\forall t \geq t_0$

$$
\begin{aligned}
\frac{dV_2(t)}{dt} &= (x - y^{(1)})^T W X^{-2} \frac{dx}{dt} \\
&= (y^{(1)} - x)^T D^{-2} DP_{AD} D\nabla f(x) \\
&\leq f(y^{(1)}) - f(x) \\
&\leq f(y^{(1)}) - f(x^{(1)}) \\
&\leq \frac{f(x^*) - f(x^{(1)})}{2} \\
&< 0,
\end{aligned}
$$

therefore $V_2(t) \to -\infty$ as $t \to +\infty$. But

$$V_2(t) = \sum_{i=1}^{n} w_i((\ln x_i) + \frac{y_i^{(1)}}{x_i}) \geq \sum_{i=1}^{n} w_i(\ln y_i^{(1)} + 1) > -\infty.$$

Hence the hypothesis is not true, and $x^{(1)}$ is an optimal solution for problem (P). □

4 Convergence of the Continuous Trajectory

Now, it comes to the key result of this paper. Theorem 4.2 below shows that if $f(x)$ is analytic, the solution of ODE system (2.1) converges to a relative interior of the optimal solution set as $t \to +\infty$. First we need the following lemma.

Lemma 4.1. *If $f(x)$ is convex and analytic, then for any two different optimal solutions x^1 and x^2 of problem (P), and any $x \in R^n$, $(x^2 - x^1)^T \nabla f(x) = 0$.*

Proof. Since $f(x)$ is convex, we have for any $\lambda \in [0,1]$,

$$\begin{aligned} f(x^1 + \lambda \Delta x) &= f(\lambda x^2 + (1-\lambda)x^1) \\ &\leq \lambda f(x^2) + (1-\lambda) f(x^1) \\ &= f(x^1) = f(x^2), \end{aligned}$$

where $\Delta x = x^2 - x^1 \neq 0$. Moreover, x^1 and x^2 are two different optimal solutions for problem (P) and $\lambda x^2 + (1-\lambda)x^1 \in \mathcal{P}^+$ for $\lambda \in [0,1]$, hence

$$f(x^1 + \lambda \Delta x) = f(x^1) = f(x^2),$$

for any $\lambda \in [0,1]$. Since $f(x)$ is analytic, then according to Corollary 1.2.5 in [20], we have for any $\lambda \in R$,

$$f(x^1 + \lambda \Delta x) = f(x^1) = f(x^2).$$

By Corollary 8.6.1 of Rockafellar [21], it follows that for any $x \in R^n$, $f(x + \lambda \Delta x)$ will be a constant function with respect to $\lambda \in R$. Hence

$$\frac{df(x + \lambda \Delta x)}{d\lambda} = (x^2 - x^1)^T \nabla f(x + \lambda \Delta x) = 0,$$

for any $\lambda \in R$. Let $\lambda = 0$ in the above equality, we have for any $x \in R^n$, $(x^2 - x^1)^T \nabla f(x) = 0$. Thus the lemma is proved. □

Theorem 4.2. *If the objective function $f(x)$ in problem (P) is analytic, then the limit set $\Omega^1(x^0)$ only contains a single point which is in the relative interior of the optimal solution set of problem (P).*

Proof. From Theorem 2.7, we know that $\Omega^1(x^0)$ is not empty. So we can choose a point $\bar{x} \in \Omega^1(x^0)$, and evidently $\bar{x} \in \mathcal{P}^+$. Without loss of generality, we assume an optimal solution x^* has the maximal number of positive components in the optimal solution set for problem (P), which is actually in the relative interior of the optimal solution set since $f(x)$ is analytic. We denote this number as k. If $k = 0$, the proof is complete, and we assume $1 \leq k \leq n$ below. Let's define

$$V_3(t) = \sum_{i=1}^{n} w_i \frac{\bar{x}_i - x_i^*}{x(t)_i}, \quad t \in [t_0, +\infty).$$

Then from Lemma 4.1 and Theorem 3.1, we have

$$\frac{dV_3(t)}{dt} = (\bar{x} - x^*)^T W X^{-2} D P_{AD} D \nabla f(x) = (\bar{x} - x^*)^T \nabla f(x) = 0,$$

so the function $V_3(t)$ is a constant. Since $\bar{x} \in \Omega^1(x^0)$, there exists a sequence $\{t_k\}$ with $t_k \to +\infty$ as $k \to +\infty$ such that $x(t_k) \to \bar{x}$ as $k \to +\infty$. Thus as $k \to +\infty$, for any index i, there are four situations:

(i) if $\bar{x}_i = x_i^* = 0$, then $\frac{\bar{x}_i - x_i^*}{x(t_k)_i} \equiv 0$;

(ii) if $\bar{x}_i = 0$, $x_i^* > 0$, then $\frac{\bar{x}_i - x_i^*}{x(t_k)_i} \to -\infty$;

(iii) if $\bar{x}_i > 0$, $x_i^* = 0$, then $\frac{\bar{x}_i - x_i^*}{x(t_k)_i} \to 1$;

(iv) if $\bar{x}_i > 0$, $x_i^* > 0$, then $\frac{\bar{x}_i - x_i^*}{x(t_k)_i} \to 1 - \frac{x_i^*}{\bar{x}_i}$.

Therefore, for any index i such that $x_i^* > 0$, $\bar{x}_i$ must be positive because $V_3(t)$ is a constant. Since x^* has the maximal number of positive components in the optimal solution set for problem (P) and $\bar{x}$ is also an optimal solution, we know $\bar{x}$ must also have k positive components and hence must be in the relative interior of the optimal solution set.

If $\bar{x}$ is not the only point in $\Omega^1(x^0)$, there must exist another point $\tilde{x} \in \Omega^1(x^0)$ with $\tilde{x} \neq \bar{x}$. Obviously, $\tilde{x}$ also has k positive components. Without loss of generality, we assume that the first k components of $\bar{x}$ and $\tilde{x}$ are positive. Then we define

$$V_4(t) = \sum_{i=1}^{n} w_i \frac{\bar{x}_i - \tilde{x}_i}{x(t)_i} = \sum_{i=1}^{k} w_i \frac{\bar{x}_i - \tilde{x}_i}{x(t)_i}, \quad t \in [t_0, +\infty).$$

Similar to $V_3(t)$, we can get that $V_4(t)$ is also a constant. Therefore if we let $x(t_k) \to \bar{x}$ and $x(t_l) \to \tilde{x}$ as $k,\ l \to +\infty$, respectively, we can get

$$\sum_{i=1}^{k} w_i (1 - \frac{\tilde{x}_i}{\bar{x}_i}) = \sum_{i=1}^{k} w_i (\frac{\bar{x}_i}{\tilde{x}_i} - 1),$$

which indicates that for $1 \leq i \leq k$, $\bar{x}_i = \tilde{x}_i$. So $\bar{x} = \tilde{x}$. Therefore, the limit set $\Omega^1(x^0)$ is a singleton. □

Acknowledgement

The authors would like to thank an anonymous reviewer for his valuable comments on an earlier version of this paper.

References

[1] I. Adler and R.D.C. Monteiro, Limiting behavior of the affine scaling continuous trajectories for linear programming problems, *Math. Program.* 50 (1991) 29–51.

[2] D.V. Anosov, S. Kh. Aranson, V.I. Arnold, I.U. Bronshtein, V.Z. Grines and Yu. S. Il'yashenko, *Ordinary Differential Equations and Smooth Dynamical Systems*, Springer, 1988.

[3] D.A. Bayer and J.C. Lagarias, The nonlinear geometry of linear programming. II Legendre transform coordinates and central trajectories, *Trans. Am. Math. Soc.* 314 (1989) 527–581.

[4] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004.

[5] E.A. Coddington and N. Levinson, *Theory of Ordinary Diffenential Equations*, McGraw-Hill Book Co., New York, 1955.

[6] I.I. Dikin, Iterative solution of problems of linear and quadratic programming, *Soviet Mathematics Doklady.* 8 (1967) 674–675 (in Russian).

[7] I.I. Dikin, On the convergence of an iterative process, *Upravlyaemye Sistemy.* 12 (1974) 54–60 (in Russian).

[8] L.M.G. Drummond and B.F. Svaiter, On well definedness of the central path, *J. Optim. Theory Appl.* 102 (1999) 223–237.

[9] A.V. Fiacco and G.P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, SIAM, Philadelphia, USA, 1990.

[10] J.C. Gilbert, C.C. Gonzaga and E. Karas, Examples of ill-behaved central paths in convex optimization, *Math. Program.* 103 (2005) 63–94.

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

[12] L.-Z. Liao, A study of the dual affine scaling continuous trajectories for linear programming, *J. Optim. Theory Appl.* 163 (2014) 548–568.

[13] J. López and H. Ramírez Cabrera, On the central paths and Cauchy trajectories in semidefinite programming, *Kybernetika.* 46 (2010) 524–535.

[14] V. Losert and E. Akin, Dynamics of games and genes: disrete versus contibuous time, *J. Math. Biology.* 17 (1983) 241–251.

[15] L. Mclinden, An analogue of Moreau's proximation theorem, with application to the nonlinear complementarity problem, *Pac. J. Math.* 88 (1980) 101–161.

[16] N. Megiddo and M. Shub, Boundary behavior of interior point algorihms for linear programming, *Math. Oper. Res.* 14 (1989) 97–146.

[17] R.D.C. Monteiro, Convergence and bounary behavior of the projective scaling trajectories for linear programmin, *Math. Oper. Res.* 16 (1991) 842–858.

[18] R.D.C. Monteiro and T. Tsuchiya, Global convergence of the affine scaling algorothm for convex quadratic programming, *SIAM J. Optim.* 8 (1998) 26–58.

[19] R.D.C. Monteiro and F. J. Zhou, On the existence and convergence of the central path for convex programming and some duality results, *Comp. Optim. Appl.* 10 (1998) 51–77.

[20] H.R. Parks and S.G. Krantz, *A primer of real analytic functions*, Birkhäuser Verlag, 1992.

[21] R.T. Rockafellar, *Convex Analysis*, Princeton Mathematical Series, No. 28 Princeton University Press, Princeton, 1970.

[22] R. Saigal, A simple proof of a primal affine scaling method, *Ann. Oper. Res.* 62 (1996) 303–324.

[23] J.J.E. Slotine and W. Li, *Applied Nonlinear Control*, Prentice Hall, New Jersey, 1991.

[24] J. Sun, A convergence proof for an affine scaling algorithm for convex quadratic programming without nondegeneracy assumptions, *Math. Program.* 60 (1993) 69–79.

[25] J. Sun, A convergence analysis for a convex version of Dikin's algorithm, *Ann. Oper. Res.* 62 (1996) 357–374.

[26] P. Tseng, I.M. Bomze and W. Schachinger, A first-order interior point method for linearly constrained smooth optimization, *Math. Program.* 127 (2011) 399–424.

[27] P. Tseng and Z.-Q. Luo, On the convergence of the affine-scaling algorithm, *Math. Program.* 56 (1992) 301–319.

[28] T. Tsuchiya, Global convergence of the affine scaling methods for degenerate linear programming problems, *Math. Program.* 52 (1991) 377–404.

[29] T. Tsuchiya, Global convergence property of the affine scaling methods for primal degenerate linear programming problems, *Math. Oper. Res.* 17 (1992) 527–557.

Manuscript received 16 June 2016
revised 8 July 2016
accepted for publication 12 July 2016

Xun Qian
Department of Mathematics, Hong Kong Baptist University
Hong Kong, P. R. China
E-mail address: 13479784@life.hkbu.edu.hk

Li-Zhi Liao
Department of Mathematics, Hong Kong Baptist University
Hong Kong, P. R. China
E-mail address: liliao@hkbu.edu.hk