Mathematical Programming 50 (1991) 29-51
North-Holland

Limiting behavior of the affine scaling continuous trajectories for linear programming problems

Ilan Adler

Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA 94720, USA

Renato D.C. Monteiro

Systems and Industrial Engineering Department, The University of Arizona, Tucson, AZ 85721, USA

Received 31 May 1988
Revised manuscript received 22 May 1989

We consider the continuous trajectories of the vector field induced by the primal affine scaling algorithm as applied to linear programming problems in standard form. By characterizing these trajectories as solutions of certain parametrized logarithmic barrier families of problems, we show that these trajectories tend to an optimal solution which in general depends on the starting point. By considering the trajectories that arise from the Lagrangian multipliers of the above mentioned logarithmic barrier families of problems, we show that the trajectories of the dual estimates associated with the affine scaling trajectories converge to the so called 'centered' optimal solution of the dual problem. We also present results related to asymptotic direction of the affine scaling trajectories. We briefly discuss how to apply our results to linear programs formulated in formats different from the standard form. Finally, we extend the results to the primal-dual affine scaling algorithm.

Key words: Interior-point methods, linear programming, Karmarkar's algorithm, logarithmic barrier function, affine scaling algorithms, continuous trajectories for linear programming.

1. Introduction

The primal affine scaling (PAS) algorithm for solving linear programming problems was first presented by Dikin [4] and later was (independently) reintroduced by Barnes [2] and Vanderbei, Meketon and Freedman [13] as a variant of the projective interior point algorithm for linear programming that was presented in the seminal work of Karmarkar [6]. The PAS algorithm is designed for linear programming problems in standard form. A similar algorithm, the so-called dual affine scaling (DAS) algorithm designed for problems in inequality form was implemented by Adler, Karmarkar, Resende and Veiga [1]. Both variants are currently the most experimented interior point algorithms for linear programming and exhibit promising results (see [1], [2] and Monma and Morton [10]).

A third variant, a primal–dual affine scaling (PDAS) algorithm was presented and analyzed by Monteiro, Adler and Resende [11]. Our main interest in this paper is to analyze the limiting behavior of the continuous trajectories of the three major variants of the affine scaling algorithms. Karmarkar [7] and Bayer and Lagarias [3] discuss some properties of the continuous PAS trajectories. Megiddo and Shub [9] analyzed the limiting behavior of these trajectories near the optimal vertex under the assumption of primal and dual non-degeneracy. Megiddo [8] presented and proved some limiting behavior of the trajectories associated with the PDAS algorithm. By presenting a weighted primal affine scaling algorithm and characterizing its infinitesimal trajectories as solutions of certain parametrized logarithmic barrier families of problems, we extend the above results and present some new results related to all three affine scaling variants with no non-degeneracy assumptions. The logarithmic barrier function method is formally studied in Fiacco and McCormick [5] in the context of nonlinear optimization.

The paper is organized as follows. In Section 2, we define a weighted PAS algorithm and present a characterization of its infinitesimal trajectories as solutions of certain parametrized logarithmic barrier families of problems. By considering the optimality conditions associated with these families of problems, we automatically obtain dual trajectories associated with the weighted PAS trajectories. We also discuss the effects of the initial conditions on the derived trajectories. In Section 3, we present the limiting behavior of the weighted PAS trajectories and its associated dual trajectories. We show in particular that, while the (weighted) PAS trajectories tend to an optimal solution of the linear program which depends on the starting point, the dual trajectories tend to the so-called (weighted) 'center' of the optimal face of the dual problem. In Section 4, we apply the results in Section 3 to show that the (weighted) dual estimates also tend to the (weighted) 'center' of the optimal face of the dual problem as x traverses a (weighted) PAS trajectory converging to an optimal solution of the linear program. In Section 5, we present some results related to the asymptotic direction of the primal and dual trajectories. In Sections 6 and 7, we briefly discuss how to apply the results of the preceding sections (2–5) to the dual affine scaling (DAS) and the primal-dual affine scaling (PDAS) trajectories, respectively. We conclude by offering some remarks in Section 8.

2. Characterization of the weighted PAS trajectories

In this section, we introduce our terminology and refresh few notions that are probably familiar to the reader. We define a family of weighted affine scaling vector fields which will be the object of our study and then characterize their trajectories as solutions of certain parametrized logarithmic barrier families of problems. This characterization will allow us to draw conclusions about the limiting behavior of these trajectories as will be described in the next section.

We start by briefly reviewing the notion of a vector field. Let $E \in \mathbb{R}^n$ be a finite dimensional affine space. Let $T(E)$ denote the subspace determined by translating the affine space E to the origin of $\mathbb{R}^n$, that is, $T(E)$ is the set of vectors tangent to the affine space E. A vector field in an open subset O of E is a mapping $\Phi : O \to T(E)$. The trajectory of the vector field Φ passing through the point $x^0 \in O$ at time $t^0 \in \mathbb{R}$

is determined by the solution curve of the following differential equation:

$$\dot{x}(t) = \Phi(x(t)), \quad x(t^0) = x^0.$$

In the following, we will be interested in a specific family of vector fields, namely, the weighted primal affine scaling vector fields. Before describing this family of vector fields, we need to introduce our terminology. Consider the linear programming problem in standard form

$$\text{(P)} \qquad \begin{aligned} \min \quad & c^{\mathrm{T}}x \\ \text{s.t.} \quad & Ax = b, \\ & x \geq 0, \end{aligned}$$

and its dual problem

$$\text{(D)} \qquad \begin{aligned} \max \quad & b^{\mathrm{T}}y \\ \text{s.t.} \quad & A^{\mathrm{T}}y + z = c, \\ & z \geq 0, \end{aligned}$$

where A is an $m \times n$ matrix and b, c are vectors of length m and n respectively. The following notation will be used throughout this paper. Let

$$S_{\mathrm{A}} = \{x \in \mathbb{R}^n;\ Ax = b\},$$

$$S_{\mathrm{F}} = \{x;\ x \in S_{\mathrm{A}},\ x \geq 0\},$$

$$S_{\mathrm{I}} = \{x;\ x \in S_{\mathrm{A}},\ x > 0\},$$

$$T_{\mathrm{A}} = \{(y, z) \in \mathbb{R}^m \times \mathbb{R}^n;\ A^{\mathrm{T}}y + z = c\},$$

$$T_{\mathrm{F}} = \{(y, z);\ (y, z) \in T_{\mathrm{A}},\ z \geq 0\},$$

$$T_{\mathrm{I}} = \{(y, z);\ (y, z) \in T_{\mathrm{A}},\ z > 0\}.$$

The sets S_{F} and T_{F} are the feasible sets of problems (P) and (D) respectively, S_{A} and T_{A} are the affine hulls of S_{F} and T_{F} respectively, and S_{I} and T_{I} (if nonempty) are the relative interiors of S_{F} and T_{F} respectively. The lower case letter e will denote the vector of all ones whose dimension is dictated by the appropriate context. If x is a lower case letter that denotes a vector $x = (x_1, \ldots, x_n)^{\mathrm{T}}$, then a capital letter will denote the diagonal matrix with the components of the vector on the diagonal, i.e., $X = \mathrm{diag}(x_1, \ldots, x_n)$. Also $\mathbb{R}^n_+$ will denote the set of n-vectors with all components strictly positive. If x and z are two n-vectors, we define their product xz to be the vector $XZe = (x_1z_1, \ldots, x_nz_n)^{\mathrm{T}}$. The inverse of x under this operation is denoted by x^{-1} and is given by $x^{-1} \equiv X^{-1}e$. In this way, expressions like x^{-1} and $x^{-2}z$ are defined if all the components of x are non-zero. No confusion should arise between the expressions xz and $x^{\mathrm{T}}z$ where the latter just denotes the inner product of x and z. Given an $m \times n$ matrix A and a subset B of the index set $\{1, \ldots, n\}$, we denote by A_B the submatrix of A associated with the index set B.

We impose the following assumptions on problem (P):

Assumption 2.1. (a) The set S_I is non-empty.

(b) The set of optimal solutions of problem (P) is non-empty and bounded.

(c) $\mathrm{rank}(A) = m$.

We should point out that assumption (c) can be discarded in our development. However, we keep it since it will simplify the arguments considerably. Assumption (b) can be verified in several alternative ways as the following proposition shows.

Proposition 2.1. *Assume that the set S_I is non-empty. Then the following conditions are equivalent*:

(a) *For all (or some)* $w = (w_1, \ldots, w_n) \in \mathbb{R}^n_+$ *and* $\mu > 0$, *the problem* $\min\{c^\mathrm{T}x - \mu \sum_{j=1}^n w_j \ln x_j;\ x \in S_\mathrm{I}\}$ *has a (unique) global solution.*

(b) *The set T_I is non-empty.*

(c) *The set of optimal solutions of problem* (P) *is non-empty and bounded.*

(d) *For all (or some)* $\bar{x} \in S_\mathrm{F}$, *the set* $\{x \in S_\mathrm{F};\ c^\mathrm{T}x \leq c^\mathrm{T}\bar{x}\}$ *is bounded.*

We refer the reader to [5] and [8] for arguments that lead to the proof of Proposition 2.1.

Next, we introduce and motivate briefly the weighted PAS algorithms. Assume that a vector $w \in \mathbb{R}^n_+$ is given. Given an interior point $\bar{x}$ for problem (P), that is, $\bar{x} \in S_\mathrm{I}$, the algorithm computes a search direction $\Delta x \equiv \Delta x(\bar{x})$ as follows. Consider the linear scaling transformation $\Psi_w : \mathbb{R}^n \to \mathbb{R}^n$ defined as $\Psi_w(x) = D^{-1}x$, $x \in \mathbb{R}^n$, where $D \equiv W^{-1/2}\bar{X}$. Observe that $\Psi_w(\bar{x}) = w^{1/2}$. In the transformed space problem (P) becomes

$$\begin{aligned} \min \quad & (Dc)^\mathrm{T} v \\ \text{s.t.} \quad & ADv = b, \\ & v \geq 0. \end{aligned}$$

The search direction d_v in the transformed space is obtained by projecting the gradient vector Dc orthogonally onto the linear subspace $\{v;\ ADv = 0\}$. Specifically, d_v is the solution of the following problem:

$$\begin{aligned} \min \quad & \tfrac{1}{2}\| Dc - v \|^2 \\ \text{s.t.} \quad & ADv = 0, \end{aligned} \tag{2.1}$$

which yields the following explicit solution:

$$d_v = [I - DA^\mathrm{T}(AD^2A^\mathrm{T})^{-1}AD]Dc,$$

together with the quantities

$$y^\mathrm{E}(\bar{x}) = (AD^2A^\mathrm{T})^{-1}AD^2c, \tag{2.2a}$$

$$z^\mathrm{E}(\bar{x}) = c - A^\mathrm{T}(AD^2A^\mathrm{T})^{-1}AD^2c, \tag{2.2b}$$

where $y^{\mathrm{E}}(\bar{x})$ is the Lagrangian multiplier associated with the constraint of (2.1) and $z^{\mathrm{E}}(\bar{x}) = c - A^{\mathrm{T}} y^{\mathrm{E}}(\bar{x})$. We shall call $y^{\mathrm{E}}(\bar{x})$ and $z^{\mathrm{E}}(\bar{x})$ the 'dual estimates' at the point $\bar{x}$ for reasons that will become apparent in Section 5. Finally, the search direction Δx in the original space is given as:

$$\Delta x = D d_v = D[I - DA^{\mathrm{T}}(AD^2A^{\mathrm{T}})^{-1}AD]Dc. \tag{2.3}$$

The next iterate $\hat{x}$ is then computed as $\hat{x} = \bar{x} - \alpha \Delta x$ where $\alpha = \alpha(\bar{x}) > 0$ is an appropriate step size which guarantees that $\hat{x} > 0$. We refer to the algorithm outlined above as the w-PAS algorithm.

Note that the w-PAS algorithm as introduced in [4], [2] and [13] uses $w \equiv e$ (that is, $D = X$). We refer to this algorithm as the PAS algorithm. Our generalization will be useful in extending the results to the primal-dual affine scaling method as will be discussed in Section 7.

We next describe a family of affine scaling vector fields which are induced by the algorithm described above. Let $E = \{x;\ Ax = b\}$ and $O \equiv S_{\mathrm{I}} \subseteq E$. Observe that $T(E) = \{x;\ Ax = 0\}$. The weighted primal affine scaling vector field with weight $w \in \mathbb{R}^n_+$, or for brevity, w-PAS vector field, is the vector field $\Phi_w : O \to T(E)$ defined as follows:

$$\Phi_w(x) \equiv D[I - DA^{\mathrm{T}}(AD^2A^{\mathrm{T}})^{-1}AD]Dc, \quad x \in O \equiv S_{\mathrm{I}},$$

where $D = W^{-1/2}X$. We are interested in studying the behavior of the trajectories of Φ_w, that is, the solution curves of the differential equation

$$\dot{q}(t) = \Phi_w(q(t)). \tag{2.4}$$

In particular, since problem (P) is in minimization form, we will be interested in the behavior of $q(t)$ as t decreases, that is, as $c^{\mathrm{T}}q(t)$ monotonically decreases. With this aim, we consider solution curves of the following reparametrized differential equation:

$$\dot{x}(\mu) = \frac{1}{\mu^2}\,\Phi_w(x(\mu)). \tag{2.5}$$

It turns out that by studying the behavior of the trajectories of (2.5), we obtain information about the trajectories of (2.4) as the following result shows.

Proposition 2.2. *Let $x^0 \in S_{\mathrm{I}}$ and $t^0 > 0$ be given. Then $q(t)$ is a solution of* (2.4) *satisfying $q(t^0) = x^0$ if and only if $x(\mu) \equiv q(t^0 + (t^0)^{-1} - \mu^{-1})$ is a solution of* (2.5) *satisfying $x(t^0) = x^0$.*

The proof of Proposition 2.2 is immediate. Note that the behavior, as t approaches $-\infty$, of a solution curve of (2.4) passing through $x^0 \in S_{\mathrm{I}}$ is completely determined by the behavior of a solution curve of (2.5) as μ approaches 0.

We now turn our efforts towards characterizing the solution curves of (2.5) as path of solutions of a logarithmic barrier family of problems. Consider the following family of problems parametrized by the penalty parameter $\mu > 0$:

$$(\mathrm{P}_\mu) \qquad \min \quad c^{\mathrm{T}}x - \mu\left[p^{\mathrm{T}}x + \sum_{j=1}^{n} w_j \ln x_j\right]$$

$$\text{s.t.} \quad Ax = b, \tag{2.6}$$

$$x > 0,$$

where $p \in \mathbb{R}^n$ is an arbitrary given vector. Since the objective function of problem (P_μ) is a strictly convex function, it follows that the global solution of (P_μ), if it exists, is completely characterized by the following Karush-Kuhn-Tucker stationary condition:

$$c - A^{\mathrm{T}}y - \mu w x^{-1} = \mu p, \tag{2.7a}$$

$$Ax = b, \quad x > 0, \tag{2.7b}$$

where $y \in \mathbb{R}^m$ is the Lagrangian multiplier associated with the equality constraint of problem (P_μ). Let $I(p)$ denote the set of parameters $\mu > 0$ such that problem (P_μ) (and hence system (2.7)) has a solution. Note that by Proposition 2.1, problem (P_μ) has a global solution if and only if the set $Y(p, \mu) \equiv \{y;\ A^{\mathrm{T}}y < c - \mu p\} \neq \emptyset$. Hence, $I(p) = \{\mu > 0;\ Y(p, \mu) \neq \emptyset\}$. The next result gives the form of the set $I(p)$.

Proposition 2.3. *The set $I(p) = (0, d_p)$ for some $d_p > 0$. Moreover, $I(0) = (0, \infty)$. Also, if the set S_{F} is bounded then $d_p = \infty$ for all $p \in \mathbb{R}^n$.*

Proof. It is straightforward to see that the set $L(p) \equiv \{\mu;\ Y(p, \mu) \neq \emptyset\}$ is convex and open. By condition (b) of Assumption 2.1 and Proposition 2.1, we have that $Y(p, 0) \neq \emptyset$. These two observations show that $L(p)$ is an open interval such that $0 \in L(p)$. Since $I(p) = L(p) \cap \mathbb{R}_+$, the first statement follows. Obviously, if $p = 0$ then $d_p = \infty$.

When the set S_{F} is bounded, it follows that the set of optimal solutions is non-empty and bounded for any objective function, and hence by Proposition 2.1, $Y(p, \mu) \neq \emptyset$, for all $\mu \in \mathbb{R}$. Thus, in this case $I(p) = (0, \infty)$, for all $p \in \mathbb{R}^n$. □

In order to simplify the notation in what follows, let $z = c - A^{\mathrm{T}}y$, for $y \in \mathbb{R}^m$. Then system (2.7) can be rewritten in an equivalent way as

$$z - \mu x^{-1}w - \mu p = 0, \tag{2.8a}$$

$$Ax - b = 0, \quad x > 0, \tag{2.8b}$$

$$A^{\mathrm{T}}y + z - c = 0. \tag{2.8c}$$

Let $(x_p(\mu), y_p(\mu), z_p(\mu))$ denote the solution of system (2.8). It can be easily verified that the path $\mu \in I(p) \to (x_p(\mu), y_p(\mu), z_p(\mu))$ has derivatives of all order. We next show that $x_p(\mu)$, $\mu \in I(p)$, is a solution of the differential equation (2.5), for all $p \in \mathbb{R}^n$. Differentiating system (2.8), we obtain

$$\dot{z}_p(\mu) + \mu x_p^{-2}(\mu)\dot{x}_p(\mu)w - x_p^{-1}(\mu)w - p = 0, \tag{2.9}$$

$$A\dot{x}_p(\mu) = 0, \tag{2.10}$$

$$A^{\mathrm{T}}\dot{y}_p(\mu) + \dot{z}_p(\mu) = 0. \tag{2.11}$$

Equations (2.8a) and (2.9) then imply that

$$\dot{z}_p(\mu) + \mu x_p^{-2}(\mu)\dot{x}_p(\mu)w = \frac{z_p(\mu)}{\mu}. \tag{2.12}$$

In order to simplify the formulas below, we drop the subscript p and the indication of variable μ. Solving for $\dot{x}$, $\dot{y}$ and $\dot{z}$ using relations (2.10), (2.11) and (2.12) and letting $D = W^{-1/2}X$, we obtain

$$\begin{aligned}
\dot{x} &= \frac{1}{\mu^2} D[I - DA^{\mathrm{T}}(AD^2A^{\mathrm{T}})^{-1}AD]Dz \\
&= \frac{1}{\mu^2} D[I - DA^{\mathrm{T}}(AD^2A^{\mathrm{T}})^{-1}AD]Dc \\
&= \frac{1}{\mu^2} \Phi_w(x), &(2.13)\\
\dot{y} &= -\frac{1}{\mu}(AD^2A^{\mathrm{T}})^{-1}AD^2z \\
&= \frac{1}{\mu}[y - (AD^2A^{\mathrm{T}})^{-1}AD^2c], &(2.14)\\
\dot{z} &= \frac{1}{\mu} A^{\mathrm{T}}(AD^2A^{\mathrm{T}})^{-1}AD^2z \\
&= \frac{1}{\mu}[A^{\mathrm{T}}(AD^2A^{\mathrm{T}})^{-1}AD^2c - A^{\mathrm{T}}y] \\
&= \frac{1}{\mu}[z + A^{\mathrm{T}}(AD^2A^{\mathrm{T}})^{-1}AD^2c - c], &(2.15)
\end{aligned}$$

where the alternative formulas above were obtained by using relation (2.8c). Expression (2.13) says that $x_p(\mu)$ is a solution of the differential equation (2.5). We summarize the above discussion in the following result.

Theorem 2.1. *Let* $(x_p(\mu), y_p(\mu), z_p(\mu))$ *denote the path of solutions for the parametrized system of equations* (2.8). *Then* $x_p(\mu)$ *is a solution of the differential equation* (2.5).

Recall that our primary interest is to study the behavior of $x_p(\mu)$ passing through some given $x^0 \in S_{\rm I}$. However, system (2.8) as developed by considering the optimal solutions $x_p(\mu)$ for the family of logarithmic barrier problems (P_μ) involves the Lagrangian multiplier $y_p(\mu)$ and its slack $z_p(\mu)$. Thus, we should also study the effect of the initialization on these trajectories. Observe that for every given $\mu_0 > 0$ and $(y^0, z^0) \in T_{\rm A}$, if we set $p = z^0/\mu_0 - w(x^0)^{-1}$ then it follows that $\mu_0 \in I(p)$ and $(x_p(\mu_0), y_p(\mu_0), z_p(\mu_0)) = (x^0, y^0, z^0)$. By Theorem 2.1, it then follows that $x_p(\mu)$ is a solution of (2.5) satisfying $x_p(\mu_0) = x^0$. In order to study the dependence on the initial conditions $\mu_0 > 0$, $x^0 \in S_{\rm I}$ and $(y^0, z^0) \in T_{\rm A}$, we denote the trajectories $x_p(\mu)$, $y_p(\mu)$ and $z_p(\mu)$ with $p = z^0/\mu_0 - w(x^0)^{-1}$ by $x(\mu; \mu_0, x^0, y^0, z^0)$, $y(\mu; \mu_0, x^0, y^0, z^0)$ and $z(\mu; \mu_0, x^0, y^0, z^0)$ respectively. We refer to $x_p(\cdot)$ as a 'weighted primal affine scaling' trajectory, or for brevity, w-PAS trajectory, and to $(y_p(\cdot), z_p(\cdot))$ as its associated 'dual trajectory'. When $w = e$, we refer to $x_p(\cdot)$ as a PAS trajectory. As should be expected, the initial condition $(y^0, z^0) \in T_{\rm A}$ does not affect the w-PAS trajectory $x(\mu; \mu_0, x^0, y^0, z^0)$. This fact is stated in the following result which also shows the behavior of the w-PAS trajectory on the initial condition $\mu_0 > 0$.

Proposition 2.4. *Given $x^0 \in S_{\rm I}$, (y^0, z^0), $(y^1, z^1) \in T_{\rm A}$ and $\mu_0, \mu_1 > 0$ then for all μ such that $0 < \mu \leq \mu_1$, we have*

$$x(\mu; \mu_1, x^0, y^1, z^1) = x(\rho(\mu); \mu_0, x^0, y^0, z^0) \tag{2.16}$$

where $\rho(\mu) \equiv (1/\mu + 1/\mu_0 - 1/\mu_1)^{-1}$.

Proof. Let $x(\mu)$ and $\tilde{x}(\mu)$ denote the expression on the left and right side of equality (2.16). By the observation following Theorem 2.1, it follows that $x(\mu)$ is a solution of (2.5) satisfying $x(\mu_1) = x^0$. On the other hand, by differentiating $\tilde{x}(\mu)$, one can easily verify that $\tilde{x}(\mu)$ is a solution of (2.5) satisfying $\tilde{x}(\mu_1) = x^0$. From the theory of differential equations, there exists a unique solution (2.5) with initial condition (x^0, μ_1). Hence, $x(\mu) = \tilde{x}(\mu)$ for all $\mu \in (0, \mu_1]$. This completes the proof of the proposition. $\square$

We next show how the initial conditions $\mu_0 > 0$ and $(y^0, z^0) \in T_{\rm A}$ affect the dual trajectories.

Proposition 2.5. *Given $x^0 \in S_{\rm I}$, (y^0, z^0), $(y^1, z^1) \in T_{\rm A}$ and $\mu_0, \mu_1 > 0$ then for all μ such that $0 < \mu \leq \mu_1$, we have*

$$y(\mu; \mu_1, x^0, y^1, z^1) = \frac{\mu}{\rho(\mu)} y(\rho(\mu); \mu_0, x^0, y^0, z^0) + \left(\frac{y^1}{\mu_1} - \frac{y^0}{\mu_0}\right)\mu, \tag{2.17}$$

$$z(\mu; \mu_1, x^0, y^1, z^1) = \frac{\mu}{\rho(\mu)} z(\rho(\mu); \mu_0, x^0, y^0, z^0) + \left(\frac{z^1}{\mu_1} - \frac{z^0}{\mu_0}\right)\mu, \tag{2.18}$$

where $\rho(\mu) \equiv (1/\mu + 1/\mu_0 - 1/\mu_1)^{-1}$.

Proof. Let $\tilde{x}(\mu)$, $\tilde{y}(\mu)$ and $\tilde{z}(\mu)$ denote the expressions in the right side of relations (2.16), (2.17) and (2.18). It can be easily shown that $(\tilde{x}(\mu), \tilde{y}(\mu), \tilde{z}(\mu))$ is a solution of system (2.8) with $p = z^1/\mu_1 - w(x^0)^{-1}$. Since such a solution is unique, the result follows. $\square$

Finally, we would like to comment about some relationship between the trajectories defined in this section and those which are discussed in [3] and [8]. Let $w \in \mathbb{R}^n_+$ be given. When $p = 0$, the path of solutions $x_0(\mu)$ of problem (P_μ) is precisely the 'weighted logarithmic barrier path' with respect to w as described in [8]. Note that $x^0 \in S_\text{I}$ lies in the path $x_0(\mu)$ if and only if for some $\mu_0 > 0$ and $y^0 \in \mathbb{R}^m$, we have $(y^0, \mu_0 w(x^0)^{-1}) \in T_\text{A}$. We shall refer to the path $x_0(\mu)$ as the 'w-central path'. In particular, when $w = e$, this path is the 'affine central path' defined in [3]. We shall refer to the e-central path as the 'central path'.

3. Limiting behavior of the w-PAS trajectories

In this section, we examine the limiting behavior of the w-PAS trajectories based on the characterization of these trajectories presented in the previous section. Given a point $x^0 \in S_\text{I}$, we show that the w-PAS trajectory through this point converges to an optimal solution of problem (P). Moreover, if the dimension of the optimal face of problem (P) is greater than zero then the limiting point is an interior point of this face that depends on the point x^0. We also show that the dual paths $(y_p(\mu), z_p(\mu))$, described in the previous section, converge to a common point, namely, the w-center of the optimal face of the dual problem, for any $p \in \mathbb{R}^n$.

We start this section by introducing the necessary terminology. Let r^* denote the common optimal value of problems (P) and (D). The optimal face of (P) (resp. (D)) is the set of points $S_\text{O} \equiv \{x;\, x \in S_\text{F},\, c^\text{T}x = r^*\}$ (resp. $T_\text{O} \equiv \{(y, z);\, (y, z) \in T_\text{F},\, b^\text{T}y = r^*\}$). The set S_O is a face of the polyhedron S_F and therefore can be expressed as the set of points $\{x \in S_\text{F};\, x_j = 0,\, j \in N\}$ for some index set $N \subseteq \{1, \ldots, n\}$. We may assume that N is the maximal set (with respect to inclusion) satisfying this property, that is, $j \in N$ if and only if $x_j = 0$ for every $x \in S_\text{O}$. Let B denote the set of indices $j \in \{1, \ldots, n\}$ such that $j \notin N$. It is well known that T_O is the face of the polyhedron T_F given by $\{(y, z) \in T_\text{F};\, z_j = 0,\, j \in B\}$ and that B is the maximal set with this property (see for example Schrijver [12]).

The next result shows that the w-PAS trajectories and its associated dual trajectories converge in objective value to the optimal value r^*.

Theorem 3.1. *Let $x^0 \in S_\text{I}$, $(y^0, z^0) \in T_\text{A}$ and $\mu_0 > 0$ be given. Consider the solution $(x(\mu), y(\mu), z(\mu)) \equiv (x_p(\mu), y_p(\mu), z_p(\mu))$ of system (2.8) with $p = z^0/\mu_0 - w(x^0)^{-1}$ and $\mu \in I(p)$. Then $\lim_{\mu \to 0} c^\text{T}x(\mu) = \lim_{\mu \to 0} b^\text{T}y(\mu) = r^*$.*

Proof. In view of Propositions 2.4 and 2.5, we may assume without loss of generality that $z^0 > 0$ and that μ_0 satisfies

$$p = z^0/\mu_0 - w(z^0)^{-1} \geq 0.$$

Since $x(\mu) > 0$ for all $\mu \in I(p)$, it then follows from (2.8a) that $z(\mu) > 0$ for all $\mu \in I(p)$. This implies that $(x(\mu), y(\mu), z(\mu)) \in S_{\mathrm{I}} \times T_{\mathrm{I}}$ for all $\mu \in I(p)$ and therefore, by the weak duality theorem of linear programming, we have that $b^{\mathrm{T}} y(\mu) \leq r^* \leq c^{\mathrm{T}} x(\mu)$ for all $\mu \in I(p)$. On the other hand, one can easily verify that if $x \in S_{\mathrm{A}}$ and $(y, z) \in T_{\mathrm{A}}$ then $c^{\mathrm{T}} x - b^{\mathrm{T}} y = x^{\mathrm{T}} z$. Therefore, $c^{\mathrm{T}} x(\mu) - b^{\mathrm{T}} y(\mu) = x(\mu)^{\mathrm{T}} z(\mu)$ for all $\mu \in I(p)$. The theorem now follows if we show that $x(\mu)^{\mathrm{T}} z(\mu)$ converges to 0 as μ tends to 0. Indeed, multiplying (2.8a) by $x(\mu)^{\mathrm{T}}$, we obtain

$$\begin{aligned} 0 \leq x(\mu)^{\mathrm{T}} z(\mu) &= \left(\sum_{j=1}^{n} w_j \right) \mu + \left(\frac{z^0}{\mu_0} - w(x^0)^{-1} \right)^{\mathrm{T}} x(\mu) \mu \\ &\leq \left(\sum_{j=1}^{n} w_j \right) \mu + \frac{\mu}{\mu_0} (z^0)^{\mathrm{T}} x(\mu) \\ &= \left(\sum_{j=1}^{n} w_j \right) \mu + \frac{\mu}{\mu_0} (c^{\mathrm{T}} x(\mu) - b^{\mathrm{T}} y^0). \end{aligned}$$

Since from Proposition 2.1, $c^{\mathrm{T}} x(\mu)$ remains bounded as μ approaches 0, the last expression implies that $x(\mu)^{\mathrm{T}} z(\mu)$ converges to 0 as μ approaches 0. This completes the proof of the theorem. $\square$

We now turn our efforts towards examining the limiting behavior of the w-PAS trajectories and its associated dual trajectories. Let $x^0 \in S_{\mathrm{I}}$, $(y^0, z^0) \in T_{\mathrm{A}}$ and $\mu_0 > 0$ be given. Set $p = z^0/\mu_0 - w(x^0)^{-1}$. In Section 2, we have seen that the path of solutions $x(\mu) \equiv x_p(\mu)$ of problem (P_μ), for $\mu \in I(p)$, is a w-PAS trajectory (Proposition 2.3 and Theorem 2.1). We are now interested in the behavior of $x(\mu)$ as μ approaches 0. The next result shows that the subvector of $x(\mu)$ corresponding to the index set N, namely, the vector $x_N(\mu)$, converges to 0 as μ approaches 0.

Lemma 3.1. (a) $\lim_{\mu \to 0} x_N(\mu) = 0$.
(b) $\lim_{\mu \to 0} c_B^{\mathrm{T}} x_B(\mu) = r^*$.

Proof. Note that by Assumption 2.1(b), Proposition 2.1 and the fact that $c^{\mathrm{T}} x(\mu)$ is decreasing in μ it follows that $x(\mu)$ lies in a compact set for all μ sufficiently small. Let $\bar{x}$ be an accumulation point of $x(\mu)$ as μ approaches 0, that is, $\bar{x} = \lim_{k \to \infty} x(\mu^k)$, where (μ^k) is a sequence of positive numbers converging to zero. By Theorem 3.1, it then follows that $c^{\mathrm{T}} \bar{x} = r^*$. Obviously $\bar{x} \in S_{\mathrm{F}}$. Hence, $\bar{x}$ lies in the optimal face of problem (P) and therefore $\bar{x}_N = 0$. Since this holds for all the accumulation points of $x(\mu)$, (a) follows. The limit in (b) is a direct consequence of (a) and Theorem 3.1. $\square$

We next analyze the limiting behavior of the subvector $x_B(\mu)$ as μ approaches 0. We first make some observations. Since $x(\mu)$ is a solution of problem (P_μ), one can easily verify that $x_B(\mu)$ is an optimal solution of the problem stated as follows.

$$(\mathrm{Q}_\mu) \qquad \begin{aligned} \max \quad & p_B^{\mathrm{T}} x_B + \sum_{j \in B} w_j \ln x_j \\ \text{s.t.} \quad & A_B x_B = b - A_N x_N(\mu), \\ & x_B > 0. \end{aligned} \tag{3.1}$$

As a consequence, for $\mu \in I(p)$, $x_B(\mu)$ satisfies the following relations which are the optimality conditions for the unique optimal solution of problem (Q_μ):

$$w_B x_B^{-1} + p_B \in H, \tag{3.2a}$$

$$A_B x_B = b - A_N x_N(\mu), \quad x_B > 0, \tag{3.2b}$$

where H is the subspace spanned by the rows of the matrix A_B.

In view of Lemma 3.1, we are led to consider the following problem which is in some sense the limit of problem (Q_μ) as μ approaches 0. Let

$$(\mathrm{Q}) \qquad \begin{aligned} \max \quad & p_B^{\mathrm{T}} x_B + \sum_{j \in B} w_j \ln x_j \\ \text{s.t.} \quad & A_B x_B = b, \\ & x_B > 0, \end{aligned} \tag{3.3}$$

and let x_B^* be the unique optimal solution of problem (Q). Obviously, $x^* \equiv (x_B^*, 0)$ is an optimal solution for (P). The next result completes the determination of the limiting behavior of the w-PAS trajectory $x(\mu)$.

Theorem 3.2. $\lim_{\mu \to 0} x_B(\mu) = x_B^*$.

Proof. Let $\bar{x}$ be an accumulation point of $x(\mu)$ as μ approaches 0. Hence, $\lim_{k \to 0} x(\mu^k) = \bar{x}$, for some sequence (μ^k) of positive numbers converging to 0. Since for all k, $x_B(\mu^k)$ satisfies system (3.2) with $\mu = \mu^k$, it follows that $\bar{x}_B$ satisfies the conditions $w_B(\bar{x}_B)^{-1} + p_B \in H$ and $A_B \bar{x}_B = b$. It turns out that these conditions are exactly the optimality conditions for the unique optimal solution x_B^* of problem (Q). Hence, $\bar{x}_B = x_B^*$ and the result follows. □

Note that the limit x_B^* of the w-PAS trajectory, as μ tends to 0, depends on $p_B = z_B^0/\mu_0 - w_B(x_B^0)^{-1}$ as can be observed by considering problem Q (of course, this dependence is realized only when the dimension of the optimal face of (P) is greater than 0). In view of Proposition 2.4, it seems that x_B^* should depend only on

x^0. Indeed, this is the case since substituting for p_B in the objective function of problem (Q) leads to

$$\begin{aligned} p_B^{\mathrm{T}} x_B &= (z_B^0/\mu_0 - w_B(x_B^0)^{-1})^{\mathrm{T}} x_B \\ &= \frac{1}{\mu_0}(c_B^{\mathrm{T}} - (y^0)^{\mathrm{T}} A_B) x_B - (w_B(x_B^0)^{-1})^{\mathrm{T}} x_B \\ &= \frac{r^* - b^{\mathrm{T}} y^0}{\mu_0} - (w_B(x_B^0)^{-1})^{\mathrm{T}} x_B \end{aligned}$$

where the last equation follows from observing the constraint of (Q) and the fact that $A_B x_B = b$ implies that $c_B^{\mathrm{T}} x_B = r^*$. Since the first term in the expression for $p_B^{\mathrm{T}} x_B$ is constant, it is clear that x_B^* is the solution of problem (Q) with $-w_B(x_B^0)^{-1}$ replacing p_B in the objective function.

In particular, Theorem 3.2 shows that the w-central path (resp. central path) converges to the so-called 'w-center' (resp. 'center') of the optimal face of (P), that is, the optimal solution of problem (Q) with $p_B = 0$. A similar result for this case was obtained in [8].

We now examine the limiting behavior of the dual trajectories associated with the w-PAS trajectories. Using Theorem 3.1 and arguments similar to the ones used in the proof of Lemma 3.1, one can show the following result.

Lemma 3.2. *Let $(y(\mu), z(\mu)) \equiv (y_p(\mu), z_p(\mu))$ be the dual trajectory associated with the w-PAS trajectory $x(\mu) = x_p(\mu)$. Then* $\lim_{\mu \to 0} z_B(\mu) = 0$.

Consider the trajectory $(y(\mu), z(\mu)) \equiv (y_p(\mu), z_p(\mu))$ associated with the w-PAS trajectory $x(\mu) = x_p(\mu)$. It follows from Section 2 that these trajectories satisfy $x(\mu) \in S_{\mathrm{I}}$, $(y(\mu), z(\mu)) \in T_{\mathrm{A}}$ and

$$x(\mu) - \mu\omega(z(\mu) - \mu p)^{-1} = 0.$$

It can be verified that these conditions are the optimality conditions for the problem stated as follows:

$$(\mathrm{D}_\mu) \qquad \begin{aligned} \max \quad & b^{\mathrm{T}} y + \mu \sum_{j=1}^{n} w_j \ln(z_j - \mu p_j) \\ \text{s.t.} \quad & A^{\mathrm{T}} y + z = c, \\ & z > \mu p. \end{aligned} \tag{3.4}$$

The point $(y(\mu), z(\mu))$ is the unique optimal solution of problem (D_μ) while $x(\mu)$ is the Lagrangian multiplier associated with the equality constraint of (D_μ). One can easily verify that $(y(\mu), z_N(\mu))$ is also an optimal solution of the problem stated as follows:

$$(\mathrm{E}_\mu) \qquad \begin{aligned} \max \quad & \sum_{j \in N} w_j \ln(z_j - \mu p_j) \\ \text{s.t.} \quad & A_N^{\mathrm{T}} y + z_N = c_N, \\ & A_B^{\mathrm{T}} y = c_B - z_B(\mu), \\ & z_N > \mu p_N. \end{aligned} \tag{3.5}$$

Now consider the following problem which arises from problem (E_μ) as μ tends to 0:

$$
\text{(E)} \qquad \begin{aligned} \max \quad & \sum_{j\in N} w_j \ln z_j \\ \text{s.t.} \quad & A_N^T y + z_N = c_N, \\ & A_B^T y = c_B, \\ & z_N > 0, \end{aligned} \tag{3.6}
$$

and let y^* and z_N^* denote its unique optimal solution. Obviously, (y^*, z^*), where $z^* = (0, z_N^*)$, is an optimal solution of problem (D). Using arguments similar to the ones used in the proof of Theorem 3.2, one can show the following result.

Theorem 3.3. *Let $(y(\mu), z(\mu)) \equiv (y_p(\mu), z_p(\mu))$ be the dual trajectory associated with the w-PAS trajectory $x(\mu) = x_p(\mu)$. Then*:

(a) $\lim_{\mu\to 0} y(\mu) = y^*$.

(b) $\lim_{\mu\to 0} z_N(\mu) = z_N^*$.

Although the limit of a w-PAS trajectory does depend on the initial condition, Theorem 3.3 shows that all dual trajectories converge to the w-center of the optimal face of the dual problem, that is, the solution of problem (E).

4. Limiting behavior of the dual estimates

Let x^k and $(y^E(x^k), z^E(x^k))$, $k = 1, 2, \ldots$, be the sequence of feasible solutions and its associated dual estimates generated by the PAS algorithm starting with $x^0 \in S_I$ (see (2.2) and (2.3)). It was shown by [2], [4] and [13] that if the linear programming problem (P) satisfies Assumption 2.1 and is primal and dual nondegenerate then x^k and $(y^E(x^k), z^E(x^k))$ converge to the unique optimal solutions x^* and (y^*, z^*) of the primal problem (P) and the dual problem (D) respectively. Actually, [4] assumes only primal non-degeneracy and shows that x^k converges to an optimal solution x^* of problem (P) satisfying $x_B^* > 0$; however, it uses smaller steps than the discrete algorithm studied in [13].

Similar results (under the same assumptions) with respect to the continuous PAS trajectories were obtained in [9].

In the previous section, we showed that a w-PAS trajectory $x(\mu)$ passing through $x^0 \in S_I$ tends to an optimal solution x^* of (P) which may depend on x^0 if problem (P) does not have a unique optimal solution. Let $x(\mu)$ denote an arbitrary w-PAS trajectory. In this section, we shall discuss the limiting behavior of the trajectory $(y^E(x(\mu)), z^E(x(\mu)))$ with no prior assumption of nondegeneracy.

Given $x \in S_{\mathrm{I}}$, recall that the dual estimate at x (see (2.2)) is given by

$$y^{\mathrm{E}}(x) = (AD^2A^{\mathrm{T}})^{-1}AD^2c, \tag{4.1a}$$

$$z^{\mathrm{E}}(x) = c - A^{\mathrm{T}}(AD^2A^{\mathrm{T}})^{-1}AD^2c, \tag{4.1b}$$

where $D = W^{-1/2}X$. It can be verified that $(y^{\mathrm{E}}(x), z^{\mathrm{E}}(x))$ satisfy the optimality conditions of the following convex quadratic programming problem where the minimization is with respect to (r, s):

$$\begin{aligned} &\min \quad \tfrac{1}{2}\|Ds\|^2 \\ &\text{s.t.} \quad A^{\mathrm{T}}r + s = c. \end{aligned} \tag{4.2}$$

Given a w-PAS trajectory $x(\mu)$, let $y^{\mathrm{E}}(\mu) \equiv y^{\mathrm{E}}(x(\mu))$ and $z^{\mathrm{E}}(\mu) \equiv z^{\mathrm{E}}(x(\mu))$ and refer to $(y^{\mathrm{E}}(\mu), z^{\mathrm{E}}(\mu))$ as the dual estimate trajectory associated with the w-PAS trajectory $x(\mu)$. In this section, our main purpose is to show that $(y^{\mathrm{E}}(\mu), z^{\mathrm{E}}(\mu))$ converges to the same limit as the dual trajectories associated with the w-PAS trajectories. Hence, as was shown in the previous section, they converge to the so-called w-center of the optimal face of problem (D).

Before stating the main result of this section, we make some observations. Note that from (2.14), (2.15) and (4.1), it follows that

$$y^{\mathrm{E}}(\mu) = y(\mu) - \mu\dot{y}(\mu), \tag{4.3}$$

$$z^{\mathrm{E}}(\mu) = z(\mu) - \mu\dot{z}(\mu). \tag{4.4}$$

Note that the quantity in the right side of (4.3) (resp. (4.4)) can be viewed as the first order Taylor approximation of the point $y(0) \equiv \lim_{\mu \to 0} y(\mu)$ (resp. $z(0) \equiv \lim_{\mu \to 0} z(\mu)$) given that we are at the point $y(\mu)$ (resp. $z(\mu)$) of the dual trajectory $y(\cdot)$ (resp. $z(\cdot)$). Expressions (4.3) and (4.4) show that, although there can be many dual trajectories $(y(\mu), z(\mu))$ (depending on the choice of the initial condition (y^0, z^0)) associated with a given w-PAS trajectory $x(\mu)$, their first order Taylor approximation as described above are the same and is equal to the dual estimate trajectory $(y^{\mathrm{E}}(\cdot), z^{\mathrm{E}}(\cdot))$.

One consequence of relations (4.3) and (4.4) is that one way to show that $y^{\mathrm{E}}(\mu)$ and $y(\mu)$ (resp. $z^{\mathrm{E}}(\mu)$ and $z(\mu)$) have the same limit as μ approaches 0 is to prove that $\dot{y}(\mu)$ (resp. $\dot{z}(\mu)$) is bounded for all μ sufficiently small. However, we have not been able to show the latter result directly. Rather, we show directly in the next theorem that $(y^{\mathrm{E}}(\mu), z^{\mathrm{E}}(\mu))$ have the same limit as $(y(\mu), z(\mu))$ when μ approaches 0. We will see in the next section that this result enables us to characterize the limiting behavior of the derivatives of the w-PAS trajectories and its associated dual trajectories.

Theorem 4.1. *Let $(x^*, y^*, z^*) = \lim_{\mu \to 0} (x(\mu), y(\mu), z(\mu))$ where $x(\mu)$ is the w-PAS trajectory passing through some $x^0 \in S_{\mathrm{I}}$ and $(y(\mu), z(\mu))$ is its associated dual trajectory. Let $(y^{\mathrm{E}}(\mu), z^{\mathrm{E}}(\mu))$ be the dual estimate trajectory associated with the w-PAS trajectory $x(\mu)$. Then $\lim_{\mu \to 0} (y^{\mathrm{E}}(\mu), z^{\mathrm{E}}(\mu)) = (y^*, z^*)$.*

Proof. Let (B, N) be the partition of the index set $\{1, \ldots, n\}$ such that $x_B^* > 0$, $x_N^* = 0$, $z_B^* = 0$ and $z_N^* > 0$. From (2.8a) we have that

$$x_N(\mu) = (z_N(\mu) - \mu p_N)^{-1} w_N \mu. \tag{4.5}$$

Since $(y^{\mathrm{E}}(x), z^{\mathrm{E}}(x))$ is the optimal solution of problem (4.2) and noting (4.5) it follows that $(y^{\mathrm{E}}(\mu), z^{\mathrm{E}}(\mu))$ is the optimal solution of the following problem where the minimization is with respect to (r, s):

$$(\mathrm{V}_\mu) \qquad \min \ \tfrac{1}{2}\|(w_B)^{-1/2} x_B(\mu) s_B\|^2 + \tfrac{1}{2}\mu^2 \|(w_N)^{1/2}(z_N(\mu) - \mu p_N)^{-1} s_N\|^2$$
$$\text{s.t.} \quad A^{\mathrm{T}} r + s = c.$$

Consider now the problem

$$(\mathrm{R}_\mu) \qquad \min \ \tfrac{1}{2}\|(w_N)^{1/2}(z_N(\mu) - \mu p_N)^{-1} s_N\|^2$$
$$\text{s.t.} \quad A_B^{\mathrm{T}} r = c_B,$$
$$A_N^{\mathrm{T}} r + s_N = c_N,$$

and let $\theta(\mu)$ and $(r(\mu), s_N(\mu))$ denote the optimal value and optimal solution of (R_μ) respectively. One can easily verify that $\theta(\mu)$ and $(r(\mu), s_N(\mu))$ converge to the optimal value and optimal solution respectively, of the following problem which arises from problem (R_μ) as μ tends to 0:

$$(\mathrm{R}) \qquad \min \ \tfrac{1}{2}\|(w_N)^{1/2}(z_N^*)^{-1} s_N\|^2$$
$$\text{s.t.} \quad A_B^{\mathrm{T}} r = c_B,$$
$$A_N^{\mathrm{T}} r + s_N = c_N.$$

Let θ^* denote the optimal value of (R). One can easily verify that (y^*, z_N^*) is the (unique) optimal solution of problem (R). Hence, $\theta^* = \tfrac{1}{2}\|(w_N)^{1/2}\|^2$. On the other hand, letting $s(\mu) = (0, s_N(\mu))$, it follows that $(r(\mu), s(\mu))$ is a feasible solution for problem (V_μ). Hence, we have

$$\mu^2 \theta(\mu) \geq \tfrac{1}{2}\|(w_B)^{-1/2} x_B(\mu) z_B^{\mathrm{E}}(\mu)\|^2 + \tfrac{1}{2}\mu^2 \|(w_N)^{1/2}(z_N(\mu) - \mu p_N)^{-1} z_N^{\mathrm{E}}(\mu)\|^2. \tag{4.6}$$

This relation obviously implies that $\lim_{\mu\to 0} (w_B)^{-1/2} x_B(\mu) z_B^{\mathrm{E}}(\mu) = 0$ and since, by Theorem 3.2, $\lim_{\mu\to 0} x_B(\mu) = x_B^* > 0$, we obtain $\lim_{\mu\to 0} z_B^{\mathrm{E}}(\mu) = 0 = z_B^*$. Since, according to Theorem 3.3, $\lim_{\mu\to 0} (z_N(\mu) - \mu p_N)^{-1} = (z_N^*)^{-1} > 0$, relation (4.6) also implies that $z_N^{\mathrm{E}}(\mu)$ is bounded for all μ sufficiently small. Since A has full row rank, it also follows that $y^{\mathrm{E}}(\mu)$ is bounded for all μ sufficiently small. Let $(\bar{y}, \bar{z})$ be an accumulation point of $(y^{\mathrm{E}}(\mu), z^{\mathrm{E}}(\mu))$ as μ tends to 0. Obviously, $\bar{z}_B = 0$. It then follows from relation (4.6) that $\theta^* \geq \tfrac{1}{2}\|(w_N)^{1/2}(z_N^*)^{-1}\bar{z}_N\|^2$. But since $(\bar{y}, \bar{z}_N)$ is feasible to (R), it follows that $(\bar{y}, \bar{z}_N)$ is an optimal solution for (R). Hence $\bar{y} = y^*$ and $\bar{z}_N = z_N^*$. Since this holds for any accumulation point of $(y^{\mathrm{E}}(\mu), z^{\mathrm{E}}(\mu))$, the result follows. □

The following corollary is a consequence of the proof of Theorem 4.1 and shows that $z_B^{\mathrm{E}}(\mu) = \mathrm{o}(\mu)$, as μ tends to 0.

Corollary 4.1. $\lim_{\mu\to 0} z_B^{\mathrm{E}}(\mu)/\mu = 0$.

Proof. Dividing expression (4.6) by $\frac{1}{2}\mu^2$ and taking the limit, we obtain

$$\theta^* \geq \tfrac{1}{2}\limsup_{\mu\to 0} \frac{1}{\mu^2}\|(w_B)^{-1/2}x_B(\mu)z_B^{\mathrm{E}}(\mu)\|^2 + \tfrac{1}{2}\|(w_N)^{1/2}\|^2$$
$$= \tfrac{1}{2}\limsup_{\mu\to 0} \frac{1}{\mu^2}\|(w_B)^{-1/2}x_B(\mu)z_B^{E}(\mu)\|^2 + \theta^*$$

which implies that

$$\lim_{\mu\to 0}\frac{1}{\mu^2}\|(w_B)^{-1/2}x_B(\mu)z_B^{\mathrm{E}}(\mu)\|^2 = 0.$$

Since $w_B > 0$ and $\lim_{\mu\to 0} x_B(\mu) > 0$, the result follows. □

5. Limiting behavior of the derivatives of the w-PAS trajectories

In this section, we analyze the limiting behavior of the derivatives of the w-PAS trajectories and their associated dual trajectories.

In order to simplify the development and the statements of the results of this section, we start by recalling the notation to be used throughout this section. Let $x^0 \in S_{\mathrm{I}}$, $(y^0, z^0) \in T_{\mathrm{A}}$ and $\mu_0 > 0$ be given. Let $p = z^0/\mu_0 - w(x^0)^{-1}$. Let $(x(\mu), y(\mu), z(\mu))$ denote the path of solutions $(x_p(\mu), y_p(\mu), z_p(\mu))$, $\mu \in I(p)$, of system (2.8). Let x^* and (y^*, z^*) denote the limits of the paths $x(\mu)$ and $(y(\mu), z(\mu))$ as μ approaches 0, respectively. Recall from Section 3 that $x_N^* = 0$, $z_B^* = 0$ and that x_B^* and (y^*, z_N^*) are the (unique) optimal solutions of problems (Q) and (E) (see (3.3) and (3.6)) respectively. Also let $(y^{\mathrm{E}}(\mu), z^{\mathrm{E}}(\mu))$ denote the dual estimate associated with $x(\mu)$, for $\mu \in I(p)$. We proved in Section 4 that the limit of $(y^{\mathrm{E}}(\mu), z^{\mathrm{E}}(\mu))$ as μ approaches 0, is equal to (y^*, z^*). As an immediate consequence of this fact, we have the following result.

Theorem 5.1. $\lim_{\mu\to 0} \dot{x}_N(\mu) = w_N(z_N^*)^{-1}$.

Proof. From (2.15) and (4.1), it follows that $z^{\mathrm{E}}(\mu) = z(\mu) - \mu\dot{z}(\mu)$. Using this relation, relations (2.12) and (2.8a), we obtain

$$\dot{x}(\mu) = w(z(\mu) - \mu\dot{z}(\mu))(z(\mu) - p\mu)^{-2}$$
$$= wz^{\mathrm{E}}(\mu)(z(\mu) - p\mu)^{-2},$$

from which the result easily follows. □

Theorem 5.1 was essentially first proved in [9] under non-degeneracy assumptions.

Some observations are in order at this point. When the dimension of the optimal face of problem (P) is 0, that is, when problem (P) has a unique optimal solution which is a vertex of S_F, the limiting behavior of $\dot{x}_B(\mu)$ is easily determined as follows. In this case, all the columns of the matrix A_B are linearly independent and we have

$$A_B\dot{x}_B(\mu)+A_N\dot{x}_N(\mu)=0$$

which implies that

$$\dot{x}_B(\mu)=-(A_B^{\mathrm{T}}A_B)^{-1}A_B^{\mathrm{T}}A_N\dot{x}_N(\mu).$$

Taking the limit in the expression above, we obtain

$$\lim_{\mu\to 0}\dot{x}_B(\mu)=-(A_B^{\mathrm{T}}A_B)^{-1}A_B^{\mathrm{T}}A_Nw_N(z_N^*)^{-1}.$$

Note that in this case the limit of $\dot{x}(\mu)$ as μ tends to 0 does not depend on the initial point $x^0\in S_1$. However, this is not necessarily the case when the dimension of the optimal face is greater than 0 as we now show.

Theorem 5.2. *The limit of $\dot{x}_B(\mu)$ as μ tends to 0 exists and is equal to the (unique) optimal solution v_B^* of the following problem where the minimization is with respect to v_B:*

$$\begin{aligned}&\min\quad \tfrac{1}{2}\|(w_B)^{1/2}(x_B^*)^{-1}v_B\|^2\\ &\text{s.t.}\quad A_Bv_B=-A_Nw_N(z_N^*)^{-1}.\end{aligned}\tag{5.1}$$

Proof. We first show that $\dot{x}_B(\mu)$ is the (unique) optimal solution of the problem

$$\begin{aligned}&\min\quad \tfrac{1}{2}\|(w_B)^{1/2}x_B^{-1}(\mu)v_B\|^2\\ &\text{s.t.}\quad A_bv_B=-A_N\dot{x}_N(\mu).\end{aligned}\tag{5.2}$$

Indeed, recall from Section 3 that since $x_B(\mu)$ is the optimal solution of problem (Q_μ) (see (3.1)), it must satisfy the conditions of system (3.2). Therefore, $\dot{x}_B(\mu)$ satisfies

$$w_Bx_B^{-2}(\mu)\dot{x}_B(\mu)\in H,\tag{5.3a}$$

$$A_B\dot{x}_B(\mu)=-A_N\dot{x}_N(\mu),\tag{5.3b}$$

or equivalently, $\dot{x}_B(\mu)$ satisfies the optimality condition for problem (5.2). Therefore, $\dot{x}_B(\mu)$ is the optimal solution of problem (5.2). Using this fact one can easily verify that $\dot{x}_B(\mu)$ is bounded as μ tends to 0. If $\bar{v}_B$ is an accumulation point of $\dot{x}_B(\mu)$ as μ approaches 0, then, by (5.3), it follows that $w_B(x_B^*)^{-2}\bar{v}_B\in H$ and $A_B\bar{v}_B=-A_Nw_N(z_N^*)^{-1}$. Hence, $\bar{v}_B$ satisfies the optimality condition for problem (5.1). This implies that $\bar{v}_B=v_B^*$. Since this holds for any accumulation point of $\dot{x}_B(\mu)$ as μ approaches 0, the result follows. □

With the aid of the previous two theorems, we now have enough information to analyze the limiting behavior of $(\dot{y}(\mu),\dot{z}(\mu))$ as μ approaches 0.

Theorem 5.3. $\lim_{\mu\to 0} \dot{z}_B(\mu) = p_B + w_B(x_B^*)^{-1}$.

Proof. Observe that by relations (2.12) and (2.8a), we have

$$\dot{z}(\mu) = p + wx^{-2}(\mu)[x(\mu) - \mu\dot{x}(\mu)]. \tag{5.4}$$

By Theorems 5.1 and 5.2, we have that $\mu\dot{x}(\mu)$ converges to 0 as μ tends to 0. From this observation, relation (5.4) and Theorem 3.2, the result follows. □

Using the previous result, one can prove the following theorem.

Theorem 5.4. *The limit of* $(\dot{y}(\mu), \dot{z}_N(\mu))$ *as* μ *tends to* 0 *exists and is equal to the (unique) optimal solution of the following problem where the minimization is with respect to* (r, s_N):

$$\begin{aligned} \min\quad & \tfrac{1}{2}\|(w_N)^{1/2}(z_N^*)^{-1}(s_N - p_N)\|^2 \\ \text{s.t.}\quad & A_B^{\mathrm{T}} r = -p_B - w_B(x_B^*)^{-1}, \\ & A_N^{\mathrm{T}} r + s_N = 0. \end{aligned} \tag{5.5}$$

The proof of Theorem 5.4 follows by noting that $(\dot{y}(\mu), \dot{z}_N(\mu))$ is the optimal solution of the following problem:

$$\begin{aligned} \min\quad & \tfrac{1}{2}\|(w_N)^{1/2}(z_N(\mu) - p_N\mu)^{-1}(s_N - p_N)\|^2 \\ \text{s.t.}\quad & A_B^{\mathrm{T}} r = -\dot{z}_B(\mu), \\ & A_N^{\mathrm{T}} r + s_N = 0, \end{aligned}$$

and by using arguments similar to the ones used in the proof of Theorem 5.2.

The next result shows that the value of the primal objective function along the w-PAS trajectory converges to the optimal value slower than the value of the dual objective function does along the dual estimate trajectory. This fact was observed in [2] and [13] with respect to the discrete version of the PAS algorithm (under non-degeneracy assumptions).

Theorem 5.5. *Let* r^* *denote the common optimal value of problems* (P) *and* (D). *Then, we have*:

(a) $c^{\mathrm{T}}x(\mu) - r^* = \mu \sum_{j\in N} w_j + \mathrm{o}(\mu)$.
(b) $r^* - b^{\mathrm{T}}y^{\mathrm{E}}(\mu) = \mathrm{o}(\mu)$.

Proof. One can easily verify that $c^{\mathrm{T}}x(\mu) - r^* = (z_N^*)^{\mathrm{T}}x_N(\mu)$ and that $r^* - b^{\mathrm{T}}y^{\mathrm{E}}(\mu) = (x_B^*)^{\mathrm{T}}z_B^{\mathrm{E}}(\mu)$. The result now follows from Corollary 4.1 and the fact that

$$\lim_{\mu\to 0} \frac{x_N(\mu)}{\mu} = \lim_{\mu\to 0} \dot{x}_N(\mu) = w_N(z_N^*)^{-1}$$

where the last equality follows from Theorem 5.1. □

6. The dual affine scaling (DAS) trajectories

In the previous sections, we restricted our attention to linear programming problems formulated in standard form. The affine scaling algorithm can be easily applied to linear programming problems formulated differently. In particular, a substantial effort was invested in exploring the same ideas as applied to the inequality form, namely, linear programming problems in the format of problem (D) of Section 2 (see [1] and [10]). Since this form is precisely the dual of a problem in standard form, it is customary to refer to the affine scaling algorithm as applied to problem (D) as the dual affine scaling (DAS) algorithm. It is obvious that all the results presented in this paper can be applied directly to the linear programming problem (D).

Specifically, consider the linear programming problem (D) of Section 2 and assume that a vector $w \in \mathbb{R}^n_+$ is given (note however that, as in the PAS algorithm, the DAS algorithm is usually defined with $w = e$ (see [1])). Given an interior point $(\bar{y}, \bar{z})$ for problem (D), that is, $(\bar{y}, \bar{z}) \in T_1$, the w-DAS algorithm computes a search direction $(\Delta y, \Delta z)$ analogous to the w-PAS algorithm (see [1]) where now the linear scaling transformation $\phi_w : \mathbb{R}^n \to \mathbb{R}^n$ defined by $\phi_w(z) = Dz, z \in \mathbb{R}^n$, with $D = W^{1/2}\bar{Z}^{-1}$, is applied to the slack z in order to transform the current slack vector $\bar{z}$ into the vector $w^{1/2}$. Similarly to the discussion on the w-PAS algorithm, problem (D) is transformed to

$$\begin{aligned} \max \quad & b^{\mathrm{T}} y \\ \text{s.t.} \quad & DA^{\mathrm{T}} y + v = Dc, \\ & v \geq 0, \end{aligned}$$

and the search direction d_y and d_v in the transformed space is then computed as the solution of the following problem:

$$\begin{aligned} \min \quad & b^{\mathrm{T}} y + \tfrac{1}{2}\|v\|^2 \\ \text{s.t.} \quad & DA^{\mathrm{T}} y + v = 0, \end{aligned}$$

which yields the explicit solution

$$d_y = -(AD^2A^{\mathrm{T}})^{-1}b, \qquad d_v = DA^{\mathrm{T}}(AD^2A^{\mathrm{T}})^{-1}b.$$

Finally, the search directions in the original space are given by

$$\Delta y = -(AD^2A^{\mathrm{T}})^{-1}b, \qquad \Delta z = A^{\mathrm{T}}(AD^2A^{\mathrm{T}})^{-1}b.$$

Observe that the quantity

$$x^{\mathrm{E}} \equiv Dd_v = D^2A^{\mathrm{T}}(AD^2A^{\mathrm{T}})^{-1}b$$

satisfies $Ax^{\mathrm{E}} = b$ and is the so-called 'primal estimate'. The next iterate $(\hat{y}, \hat{z})$ is computed as $\hat{y} = \bar{y} - \alpha\Delta y$ and $\hat{z} = \bar{z} - \alpha\Delta z$ where $\alpha \equiv \alpha(\bar{y}, \bar{z}) > 0$ is an appropriate step size which guarantees that $\hat{z} > 0$. In a manner similar to Section 2, we are interested in studying the trajectories induced by the vector field $V_w: T_I \times \mathbb{R} \to \{(u, v);\ A^Tu + v = 0\}$ defined as

$$V_w((y, z), \mu) = \frac{1}{\mu^2}(\Delta y, \Delta z).$$

Thus similar to the w-PAS case the trajectories $(y(\mu), z(\mu))$ are characterized as the solutions of the following logarithmic barrier family of problems:

$$\begin{aligned} \max \quad & b^Ty + \mu\left[p^Tz + \sum_{j=1}^n w_j \ln z_j\right] \\ \text{s.t.} \quad & A^Ty + z = c, \\ & z > 0, \end{aligned}$$

which leads to the following relations between the trajectories $(y(\mu), z(\mu))$ and its associated Lagrangian 'primal trajectories' $x(\mu)$:

$$x(\mu) - \mu wz^{-1}(\mu) = p\mu$$

where $p = x^0/\mu^0 - w(z^0)^{-1}$, x^0 is an arbitrary vector satisfying $Ax^0 = b$ and $\mu_0 > 0$. Similarly, one can show that $x^E(\mu) = x(\mu) - \mu\dot{x}(\mu)$. Then, all the results in the previous sections will apply to these constructs by simply 'dualizing' in the sense that x and (y, z) are interchanged. This task is rather trivial so we leave out the details.

7. Analysis of the primal–dual affine scaling (PDAS) trajectories

In this section, we apply all the results obtained in Sections 2 to 5 to the trajectories generated by the primal-dual affine scaling algorithm.

The primal-dual affine scaling algorithm has been presented and analyzed in [11]. Here, we briefly review the directions generated by the PDAS algorithm. Given a point $v = (x, y, z) \in S_I \times T_I$, the PDAS algorithm computes a search direction $\Delta v \equiv \Delta v(v) \equiv (\Delta x, \Delta y, \Delta z) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ which is a solution of the following system of linear equations:

$$\begin{aligned} Z\Delta x + X\Delta z &= xz, \\ A\Delta x &= 0, \\ A^T\Delta y + \Delta z &= 0. \end{aligned} \tag{7.1}$$

This system yields the following explicit expression for Δx, Δy and Δz:

$$\begin{aligned} \Delta x &= [Z^{-1} - Z^{-1}XA^T(AZ^{-1}XA^T)^{-1}AZ^{-1}]xz, \\ \Delta y &= -[(AZ^{-1}XA^T)^{-1}AZ^{-1}]xz, \\ \Delta z &= [A^T(AZ^{-1}XA^T)^{-1}AZ^{-1}]xz. \end{aligned}$$

Letting $D=(Z^{-1}X)^{1/2}$ and using the fact that $A^{\mathrm{T}}y+z=c$ and $Ax=b$, we obtain

$$\Delta x = D[I - DA^{\mathrm{T}}(AD^2A^{\mathrm{T}})^{-1}AD]Dc,$$

$$\Delta y = -(AD^2A^{\mathrm{T}})^{-1}b,$$

$$\Delta z = A^{\mathrm{T}}(AD^2A^{\mathrm{T}})^{-1}.$$

We are interested in obtaining results about the limiting behavior of the trajectories induced by the vector field $\Delta v(v)$, $v \equiv (x, y, z) \in S_{\mathrm{I}} \times T_{\mathrm{I}}$. With this aim, we consider the solution curves of the following reparametrized differential equation:

$$\dot{v}(\mu) = \frac{1}{\mu}\,\Delta v(v(\mu)), \quad v(\mu_0) = v^0, \tag{7.2}$$

where $v(\mu) \equiv (x(\mu), y(\mu), z(\mu)) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ and where $\mu_0 > 0$ and $v^0 \equiv (x^0, y^0, z^0) \in S_{\mathrm{I}} \times T_{\mathrm{I}}$ are the initial conditions. Equivalently, in view of the definition of $\Delta v(v)$ given by system (7.1), the differential equation (7.2) can be rewritten as

$$z(\mu)\dot{x}(\mu) + x(\mu)\dot{z}(\mu) = \frac{x(\mu)z(\mu)}{\mu}, \tag{7.3a}$$

$$A\dot{x}(\mu) = 0, \tag{7.3b}$$

$$A^T\dot{y}(\mu) + \dot{z}(\mu)0, \tag{7.3c}$$

$$x(\mu_0) = x^0, \quad y(\mu_0) = y^0, \quad z(\mu_0) = z^0. \tag{7.3d}$$

The following result characterizes the solution of (7.3) and is an immediate consequence of the results of Section 2.

Theorem 7.1. *Let $v(\mu) \equiv (x(\mu), y(\mu), z(\mu))$ be the (unique) solution of system* (2.8) *with $p \equiv 0$ and $w \equiv x^0z^0/\mu_0$. Then $v(\mu)$ is the (unique) solution of* (7.3). *As a consequence, it follows that $x(\mu)$ is the (unique) optimal solution of*

$$\begin{aligned} \min \quad & c^{\mathrm{T}}x - \mu \sum_{j=1}^{n} w_j \ln x_j \\ \text{s.t.} \quad & Ax = b, \\ & x > 0, \end{aligned} \tag{7.4}$$

while $(y(\mu), z(\mu))$ is the (unique) optimal solution of

$$\begin{aligned} \max \quad & b^{\mathrm{T}}y + \mu \sum_{j=1}^{n} w_j \ln z_j \\ \text{s.t.} \quad & A^{\mathrm{T}}y + z = c, \\ & z > 0. \end{aligned} \tag{7.5}$$

Proof. It follows from (2.10), (2.11), (2.12), (2.8a) and the fact that $p = 0$ that $v(\mu)$ satisfies (7.3a), (7.3b) and (7.3c). Also, since $v^0 \in S_I \times T_I$, $p = 0$ and $w = x^0 z^0/\mu_0$, it follows that v^0 is a solution of system (2.8) with $\mu = \mu_0$. Hence $v(\mu_0) = v^0$. This shows that $v(\mu)$ is a solution of (7.3). To show that $x(\mu)$ and $(y(\mu), z(\mu))$ are the optimal solutions of problems (P) and (D) respectively, just note that (2.8) with $p = 0$ is the optimality condition for both problems. □

For a detailed discussion of the relationship between problems (7.4) and (7.5), see also [8].

Theorem 7.1 shows that the family of PDAS trajectories are precisely the w-central paths as defined in Section 2. Thus, all the results obtained in Sections 2 to 5 are applicable to these trajectories as well. We summarize these results in the following two theorems.

Theorem 7.2. *Let $x^0 \in S_I$ and $(y^0, z^0) \in T_I$ be given. Let $(x(\mu), y(\mu), z(\mu))$ be the PDAS trajectory passing through (x^0, y^0, z^0). Let $w = x^0 y^0$. Then*

$$\lim_{\mu \to 0} (x(\mu), y(\mu), z(\mu)) = (x^*, y^*, z^*)$$

where x^* and (y^*, z^*) are the w-center points of the optimal faces of the primal (P) *and the dual* (D) *problems respectively.*

It should be noted that Theorem 7.2 was first proved in [8].

Theorem 7.3. *Let $x^0 \in S_I$ and $(y^0, z^0) \in T_I$ be given. Let $(x(\mu), y(\mu), z(\mu))$ be the PDAS trajectory passing through (x^0, y^0, z^0). Let $w = x^0 z^0$. Let x^* and (y^*, z^*) denote the w-center points of the optimal faces of the primal* (P) *and the dual* (D) *problems respectively. Then, the limit of $(\dot{x}(\mu), \dot{y}(\mu), \dot{z}(\mu))$, as μ tends to 0, exists and its value is as follows:*

(a) $\lim_{\mu \to 0} \dot{x}_N(\mu) = w_N (z_N^*)^{-1}$ *and* $\lim_{\mu \to 0} \dot{x}_B(\mu)$ *is equal to the unique optimal solution of problem (5.1).*

(b) $\lim_{\mu \to 0} \dot{z}_B(\mu) = w_B (x_B^*)^{-1}$ *and* $\lim_{\mu \to 0} \dot{z}_N(\mu)$ *is equal to the unique optimal solution of problem (5.5) with p_B and p_N replaced by* 0.

8. Remarks

It is still an open question whether the actual discrete PAS (or DAS) algorithm converges to an optimal solution without assuming both primal and dual non-degeneracy. We believe that the results presented here are a step towards proving that all the accumulation points (and possibly the limit) of the sequence of iterates (x^k) generated by the PAS algorithm (with appropriate step size) lies in the optimal face of the primal problem.

A closely related problem to the one mentioned above is the convergence of the dual estimate in the discrete PAS (or DAS) algorithm. We again believe that the results presented in this paper provide some insight to the observed fact (see e.g. [1] and [10]) that the 'dual estimates' (or 'primal estimates' in the case of the DAS algorithm) do converge to an optimal solution of the dual (primal) problem.

In contrast to the discrete PAS (or DAS) algorithm, it has been proved in [11] that the PDAS algorithm converges (in polynomial time) to an optimal and dual solution without any non-dengeracy assumption. We believe that the inherently simpler structure of the PDAS trajectories (the w-central trajectories) may provide some insight to the polynomial convergence of the PDAS algorithm.

Acknowledgement

This research was partially funded by the United States Navy Office of Naval Research under contract N00014-87-K-0202 and the Brazilian Post-Graduate Education Federal Agency (CAPES). Their financial support is gratefully acknowledged.

References

[1] I. Adler, N. Karmarkar, M.G.C. Resende and G. Veiga, "An implementation of Karmarkar's algorithm for linear programming," *Mathematical Programming* 44 (1989) 297-335.

[2] E.R. Barnes, "A variation on Karmarkar's algorithm for solving linear programming problems," *Mathematical Programming* 36 (1986) 174-182.

[3] D.A. Bayer and J.C. Lagarias, "The nonlinear geometry of linear programming," *Transactions of the American Mathematical Society* 314 (1989) 499-581.

[4] I.I. Dikin, "Iterative solution of problems of linear and quadratic programming," *Soviet Mathematics Doklady* 8 (1967) 674-675.

[5] A. Fiacco and G. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques* (John Wiley and Sons, New York, 1955).

[6] N. Karmarkar, "A new polynomial time algorithm for linear programming," *Combinatorica 4* (1984) 373-395.

[7] N. Karmarkar, Talk at the University of California at Berkeley (Berkeley, CA, 1984).

[8] N. Megiddo, "Pathways to the optimal set in linear programming," in: N. Megiddo, ed., *Progress in Mathematical Programming: Interior-Point Algorithms and Related Methods* (Springer, Berlin, 1989) pp. 131-158.

[9] N. Megiddo and M. Shub, "Boundary behavior of interior point algorithms for linear programming," *Mathematics of Operations Research* 14 (1989) 97-146.

[10] C.L. Monma and A.J. Morton, "Computational experience with a dual affine variant of Karmarkar's method for linear programming," *Operations Research Letters* 6 (1987) 261-267.

[11] R.D.C. Monteiro, I. Adler and M.G.C. Resende, "A polynomial-time primal-dual affine scaling algorithm for linear and convex quadratic programming and its power series extension," *Mathematics of Operations Research* 15 (1990) 191-214.

[12] A. Schrijver, *Theory of Linear and Integer Programming* (John Wiley & Sons, New York 1986).

[13] R.J. Vanderbei, M.S. Meketon and B.A. Freedman, "A modification of Karmarkar's linear programming algorithm," *Algorithmica* 1 (1986) 395-407.