Mathematical Programming 67 (1994) 169–187

# Asymptotic analysis of the exponential penalty trajectory in linear programming

R. Cominetti, J. San Martín *

*Departamento de Ingeniería Matemática, Universidad de Chile, Casilla 170/3 Correo 3, Santiago, Chile*

Received 21 December 1992; revised manuscript received 4 February 1994

**Abstract**

We consider the linear program $\min\{c'x: Ax \leqslant b\}$ and the associated exponential penalty function $f_r(x) = c'x + r\Sigma\exp[(A_ix - b_i)/r]$. For $r$ close to 0, the unconstrained minimizer $x(r)$ of $f_r$ admits an asymptotic expansion of the form $x(r) = x^* + rd^* + \eta(r)$ where $x^*$ is a particular optimal solution of the linear program and the error term $\eta(r)$ has an exponentially fast decay. Using duality theory we exhibit an associated dual trajectory $\lambda(r)$ which converges exponentially fast to a particular dual optimal solution. These results are completed by an asymptotic analysis when $r$ tends to $\infty$: the primal trajectory has an asymptotic ray and the dual trajectory converges to an interior dual feasible solution.

*AMS Subject Classification*: 90C05, 90C25, 90C31, 49M30

*Keywords*: Linear programming; Exponential penalty; Optimal trajectory; Asymptotic expansion; Interior point methods

## 1. Introduction

Polynomial methods for linear programming (LP) have become a central theme in mathematical programming since Khachiyan [13] proved the polynomial solvability of LP. Khachiyan's result solved an important problem in complexity theory, but failed to provide a competitive algorithm which could challenge the (non-polynomial but practically efficient) Simplex method.

A most important step towards this end was achieved by Karmarkar [12] who proposed an algorithm based on potential functions which, besides solving LP in polynomial time, proved to be efficient in practice. The close relation between Karmarkar's potential function

* Corresponding author. Both authors partially supported by FONDECYT.


*SSDI* 0025-5610(94)00024-N

and the classical logarithmic barrier function of Frisch opened the way to further progress. Since then, the number of contributions and new polynomial methods for LP has grown rapidly. We refer to the recent review [9] for a detailed and unified account of the main developments in the field.

These achievements have definitely brought LP into the realm of nonlinear programming but, with few exceptions [5, 11, 16], the attention has been primarily directed towards the methods and properties related to the logarithmic barrier function. It is the purpose of this work to investigate an alternative tool from nonlinear programming, the *exponential penalty function*, in the context of linear programming.

To be more precise, let us consider the linear program

$$\text{(LP)} \qquad \min_x \{c'x: Ax \leqslant b\}$$

and its unconstrained penalized version

$$(\mathrm{P}_r) \qquad \min_x c'x + r \sum \exp[(A_i x - b_i)/r] .$$

Under very mild assumptions problem $(\mathrm{P}_r)$ has a unique solution $x(r)$. Since the problems $(\mathrm{P}_r)$ *tend* to (LP) as $r$ goes to 0, a first goal in our study is to investigate the asymptotic behavior of the trajectory $x(r)$ and its relation with the solution set of (LP). We shall prove that, essentially, the trajectory $x(r)$ is a straight line directed towards the center of the optimal face of (LP), namely

$$x(r) = x^* + rd^* + \eta(r)$$

where $x^*$ is a (central) optimal solution of (LP) which we call *centroid*, the directional derivative $d^*$ is completely characterized in variational terms, and the error $\eta(r)$ goes to zero exponentially fast, that is, at the speed of $\exp(-\mu/r)$ for some $\mu > 0$.

Our second goal is to investigate the asymptotic behavior of the dual trajectory $\lambda(r)$, defined as the unique solution of the following dual of $(\mathrm{P}_r)$,

$$(\mathrm{D}_r) \qquad \min_\lambda \left\{ b'\lambda + r \sum \lambda_i (\ln \lambda_i - 1) : A'\lambda + c = 0,\ \lambda \geqslant 0 \right\} .$$

We remark that the classical linear programming dual of (LP) is

$$\text{(D)} \qquad \min_\lambda \{b'\lambda: A'\lambda + c = 0,\ \lambda \geqslant 0\} ,$$

so we may interpret $(\mathrm{D}_r)$ as a penalty method which introduces the positivity constraints of (D) into the objective function through the penalty term "$r\Sigma\lambda_i(\ln \lambda_i - 1)$". The dual trajectory is closely related to the primal one as $\lambda_i(r) = \exp[(A_i x(r) - b_i)/r]$. We prove that this trajectory is essentially a constant

$$\lambda(r) = \lambda^* + \nu(r) ,$$

where $\lambda^*$ is a center of the optimal face of (D) and the error $\nu(r)$ goes to zero exponentially fast.

The previous results are completed by an asymptotic analysis when $r$ tends to $\infty$. For the primal trajectory we find that it has an asymptotic ray, namely,

$$x(r) = x^\infty + rd^\infty + \rho(r)$$

with the error term $\rho(r)$ tending to 0 when $r \uparrow \infty$, while the dual trajectory converges towards an interior point of the dual feasible polytope (see Fig. 1).

Similar limiting properties for the optimal trajectory associated with the logarithmic barrier function, the so-called *central path*, have been extensively studied in recent years (see [10, 14] and references therein).

The "almost straight" asymptotic character of the exponential trajectory (and our limited computational experience) suggests that a path-following method should have no trouble to follow this trajectory. More precisely, we suggest to approximately trace the primal path $x(r)$, which solves an unconstrained strictly convex problem and converges to $x^*$ at a *reasonable* linear speed, and to check for convergence by looking at the dual path which tends to $\lambda^*$ much faster. While there exist several methods based on the idea of exponential penalties [1, 15, 19], the asymptotic analysis presented in this paper is more directly connected with predictor–corrector methods as the one studied in [4] in the setting of nonlinear programming.

Another favorable fact is that, in contrast to the interior point methods, the exponential penalty is everywhere defined and does not need strictly interior points in the primal problem, which is certainly a desirable feature. However, our analysis here is just asymptotic and it is not clear whether a path-following method based on exponential penalties will yield a polynomial method for LP.

The paper is organized as follows. In Section 2 we settle the notation and we present the primal and dual trajectories. In Section 3 we define the *centroid* and prove the convergence of the primal trajectory towards this particular optimal solution of (LP), while in Section 4 we establish the convergence of the dual trajectory towards a dual optimal solution.

The results concerning the asymptotic behavior of the trajectories when $r$ tends to 0,

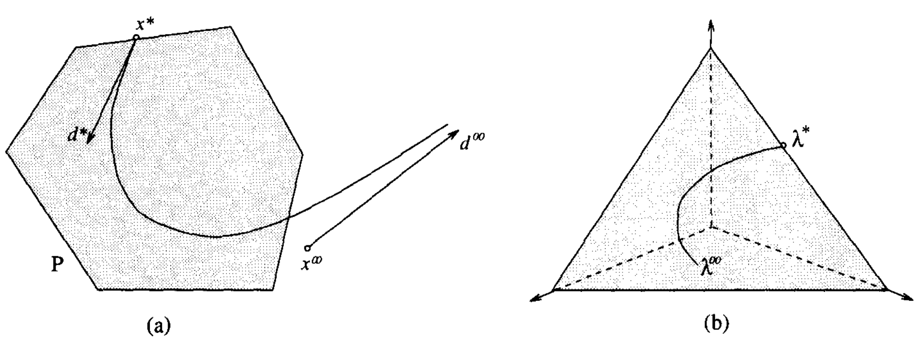


Fig. 1. A two dimensional example of primal (a) and dual (b) trajectories.

including the exponential decay of the error terms, are proved in Section 5. These expansions imply that all higher order derivatives of the trajectories at $r=0$ must vanish as soon as they exist, thus in Section 6 we point out a particular instance where $C^\infty$ differentiability of the trajectories at $r=0$ can be ensured.

The asymptotic analysis of the primal and dual trajectories when $r$ tends to $\infty$ is presented in Section 7.

We shall assume some familiarity with convex analysis, and particularly with the abstract theory of convex duality for which we refer to [17]. While this is not strictly necessary since all the results in this paper may be proved by direct arguments, duality theory allows shorter proofs and somewhat clarifies which properties are general facts and which are specific to the exponential penalty.

## 2. The primal and dual trajectories

Let us consider the linear program

$$\text{(LP)} \qquad \alpha := \min_x \{c'x \colon Ax \leqslant b\}$$

where $c \in \mathbb{R}^n$, $A$ is an $m \times n$ matrix of full rank $n \leqslant m$, and $b \in \mathbb{R}^m$. We denote by $A_i$ the rows of $A$ for $i \in I := \{1, \ldots, m\}$, and $P := \{x \in \mathbb{R}^n \colon A_i x \leqslant b_i,\ i \in I\}$ the feasible polytope of (LP). We assume throughout that (LP) has a *nonempty and bounded* optimal solution set

$$S_0 := \{x \in \mathbb{R}^n \colon Ax \leqslant b,\ c'x = \alpha\} .$$

This boundedness assumption is equivalent (see e.g. [18]) to the existence of a strictly positive feasible solution for the dual problem (D), that is, a certain $\lambda \in \mathbb{R}^m$ such that $c + A'\lambda = 0$ and $\lambda_i > 0$ for all $i = 1, \ldots, m$.

We associate with (LP) the unconstrained problem

$$(\mathrm{P}_r) \qquad \min_x f_r(x)$$

where $f_r$ is the exponential penalty function

$$f_r(x) := c'x + r \sum_{i \in I} \exp[(A_i x - b_i)/r] .$$

Using convex duality theory [17], we construct a dual problem for $(\mathrm{P}_r)$ by considering the perturbation function $\varphi_r \colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ defined as

$$\varphi_r(x, u) = c'x + r \sum_{i \in I} \exp[(A_i x - b_i + u_i)/r] .$$

The dual problem, which we shall denote $(\mathrm{D}_r)$, seeks to minimize the Fenchel conjugate $v_r^*$ of the marginal function $v_r(u) = \inf_x \varphi_r(x, u)$, and can be explicited by using the identity $v_r^*(\lambda) = \varphi_r^*(0, \lambda)$. Indeed, a straightforward computation gives for all $(y, \lambda) \in \mathbb{R}^n \times \mathbb{R}^m$,

$$\varphi_r^*(y,\lambda)=\begin{cases} b'\lambda+r\Sigma_{i\in I}\lambda_i(\ln\lambda_i-1) & \text{if } A'\lambda+c=y,\ \lambda\geqslant 0\,,\\ +\infty & \text{otherwise}\,,\end{cases}$$

so that we get

$$(\mathrm{D}_r)\qquad \min_{\lambda}\left\{b'\lambda+r\sum_{i\in I}\lambda_i(\ln\lambda_i-1):A'\lambda+c=0,\ \lambda\geqslant 0\right\}.$$

Conversely, taking $\varphi_r^*$ as a perturbation function for $(\mathrm{D}_r)$ and $w_r(y)=\inf_\lambda\varphi_r^*(y,\lambda)$ its associated (dual) marginal function, the dual of $(\mathrm{D}_r)$ turns out to be precisely the primal problem $(\mathrm{P}_r)$. In other words, problem $(\mathrm{P}_r)$ corresponds to the minimization of $w_r^*$. For further details on the general theory of convex duality, the reader may consult [17, pp. 18–21].

Incidentally, we observe that problem $(\mathrm{D}_r)$ falls into the framework of entropy minimization problems, for which several specific solution methods have been proposed (see e.g. [2, 3, 6, 7] and references therein). However, in our case the entropy term has no special meaning from the statistical mechanics or information theory point of view, and we just interpret it as a penalty term associated with the constraint $\lambda\geqslant 0$.

In what follows we shall say that an optimization problem is *coercive* if the level sets of the function to be minimized are bounded. This property not only yields the existence of optimal solutions but, from an algorithmic point of view, it ensures the boundedness of the minimizing sequences generated by any descent method.

**Proposition 2.1.** *Problems* $(\mathrm{P}_r)$ *and* $(\mathrm{D}_r)$ *are strictly convex and coercive. The optimal values of these problems satisfy* $v(\mathrm{P}_r)+v(\mathrm{D}_r)=0$, *and the corresponding unique solutions* $x(r)$ *and* $\lambda(r)$ *are related by*

$$\lambda_i(r)=\exp[(A_ix(r)-b_i)/r]\ .$$

**Proof.** Let $v_r$ and $w_r$ be the primal and dual marginal functions introduced above. Clearly $\operatorname{dom} v_r=\mathbb{R}^m$ so that [17, Theorem 10 (a)] implies that $v_r^*$ has bounded level sets and then $(\mathrm{D}_r)$ is coercive. Similarly, we have $\operatorname{dom} w_r=c+A'\mathbb{R}_+^m$ and since (D) has a strictly positive feasible solution and $A'$ is surjective, we get $0\in\operatorname{int}(\operatorname{dom} w_r)$. Using [17, Theorem 10 (a)] once again it follows that $w_r^*$ has bounded level sets, hence $(\mathrm{P}_r)$ is coercive.

The strict convexity of $(\mathrm{D}_r)$ is an immediate consequence of the strict convexity of the function $t\to t\ln t$. On the other hand, we have $\nabla^2 f_r(x)=(1/r)A'DA$ where $D:=\operatorname{diag}\{\exp[(A_ix-b_i)/r]:i\in I\}$ and, since $A$ has full rank, this Hessian is positive definite so that $f_r$ is strictly convex.

The equality $v(\mathrm{P}_r)+v(\mathrm{D}_r)=0$ and the link between the primal and dual optimal solutions follow from [17, Theorem 18 (b)] and [17, Theorem 15 (e) $\Leftrightarrow$ (f)] respectively. $\square$

The solution $x(r)$ of $(\mathrm{P}_r)$ is also characterized by the stationarity condition $\nabla f_r(x(r))=0$

and, since $\nabla^2 f_r(x(r))$ is positive definite, the implicit function theorem tells us that the trajectory $x(r)$, and a fortiori $\lambda(r)$, are $C^\infty$ on $(0, +\infty)$. The derivative of $x(r)$ satisfies

$$\nabla^2 f_r(x(r))\dot{x}(r) + \frac{\partial[\nabla f_r]}{\partial r}(x(r)) = 0\,,$$

that is,

$$(A'DA)\dot{x}(r) = \sum_{i\in I} \lambda_i(r)\ln\lambda_i(r)A_i' = A'D\ln\lambda(r)\,, \tag{1}$$

where we denote $\ln\lambda(r) := (\ln\lambda_1(r), \ldots, \ln\lambda_m(r))'$ and $D := \operatorname{diag}\{\lambda_i(r) : i\in I\}$.

## 3. Convergence of the primal trajectory

In the next propositions we study the limit of $x(r)$ as $r$ tends to 0.

**Proposition 3.1.** *The primal trajectory $x(r)$ stays bounded as $r$ approaches zero and every accumulation point of $x(r)$ is an optimal solution for* (LP).

**Proof.** Let $x\in S_0$ be an optimal solution of (LP) so that

$$c'x(r) + r\sum_{i\in I}\exp[(A_ix(r) - b_i)/r] \leqslant c'x + rm = \alpha + rm\,. \tag{2}$$

Suppose $\|x_k\| \to \infty$ where $x_k := x(r_k)$ for a given sequence $r_k \downarrow 0$, and assume with no loss of generality that $x_k/\|x_k\| \to d$ for some $d \neq 0$. From (2) we obtain $A_ix_k \leqslant b_i + r_k\ln[(\alpha + r_km - c'x_k)/r_k]$, so that for some constant $K$ and $k$ large

$$A_i\frac{x_k}{\|x_k\|} \leqslant \frac{b_i}{\|x_k\|} + \frac{r_k}{\|x_k\|}\ln\left(\frac{K\|x_k\|}{r_k}\right),$$

from which we deduce $A_id \leqslant 0$. Also, (2) implies $c'x_k \leqslant \alpha + r_km$ so that $c'd \leqslant 0$, which combined with $Ad \leqslant 0$ contradicts the boundedness of $S_0$.

This contradiction proves that $x(r)$ stays bounded as $r$ tends to zero. Then, using (2) we may find a constant $M$ such that for all small $r > 0$

$$r\sum_{i\in I}\exp[(A_ix(r) - b_i)/r] \leqslant M\,.$$

It follows that $\limsup_{r\downarrow 0} A_ix(r) \leqslant b_i$, proving that every accumulation point of $x(r)$ must be feasible for (LP). Similarly, from (2) we get $c'x(r) \leqslant \alpha + rm$, so that $\limsup_{r\downarrow 0} c'x(r) \leqslant \alpha$, and the accumulation points of $x(r)$ are not only feasible but optimal for (LP). □

We shall prove that in fact the curve $x(r)$ converges towards a particular optimal solution of (LP) which is described next.

Let $I_0 := \{i \in I: A_i x = b_i \text{ for all } x \in S_0\}$ be the set of constraints saturated in $S_0$. Define $\phi_0(x) := \min\{b_i - A_i x: i \notin I_0\}$ and let $S_1$ be the optimal solution set of

$$t_1 := \max\{\phi_0(x): x \in S_0\} .$$

Geometrically, $t_1$ and $S_1$ are found by considering the polytopes $P^t := \{x \in S_0: A_i x \leqslant b_i - t, i \notin I_0\}$ and taking the largest $t$ so that $P^t$ is nonempty (see Fig. 2). Thus, we are *pushing in* all non-saturated constraints until some of them become saturated, namely, those in $J_1 := \{i \notin I_0: A_i x = b_i - t_1 \text{ for all } x \in S_1\}$.

Letting $I_1 := I_0 \cup J_1$, $\phi_1(x) := \min\{b_i - A_i x: i \notin I_1\}$, and $S_2$ the optimal set of

$$t_2 := \max\{\phi_1(x): x \in S_1\} ,$$

we may define $J_2 := \{i \notin I_1: A_i x = b_i - t_2 \text{ for all } x \in S_2\}$, $I_2 := I_1 \cup J_2$ and $\phi_2(x) := \min\{b_i - A_i x: i \notin I_2\}$, in order to find $t_3$, $S_3$ and so on.

Since the sets $J_j$ are non-empty, the sequence $I_0 \subset I_1 \subset \cdots \subset I_k$ is strictly increasing and after finitely many steps we have $I_k = I$. The corresponding decreasing sequence of polytopes $S_0 \supset S_1 \supset \cdots \supset S_k$ satisfies

$$S_j = \left\{ x: \begin{array}{l} A_i x = b_i,\ i \in I_0 \\ A_i x = b_i - t_1,\ i \in J_1;\ \ldots;\ A_i x = b_i - t_j,\ i \in J_j \\ A_i x \leqslant b_i - t_j,\ i \notin I_j \end{array} \right\} .$$

In particular $S_k$ is defined only by equalities and, since $A$ has full rank, it is reduced to a singleton $\{x^*\}$.

**Definition 1.** The point $x^*$ defined by the process above will be called the *centroid* of the optimal face $S_0$.

We remark that

$$0 < t_1 < t_2 < \cdots < t_k , \tag{3}$$

and that $S_j$ is also characterized by the following set of inequalities,

$$S_j = \left\{ x: \begin{array}{l} c'x \leqslant \alpha;\ A_i x \leqslant b_i,\ i \in I_0 \\ A_i x \leqslant b_i - t_1,\ i \in J_1;\ \ldots;\ A_i x \leqslant b_i - t_j,\ i \in J_j \\ A_i x \leqslant b_i - t_j,\ i \notin I_j \end{array} \right\} . \tag{4}$$

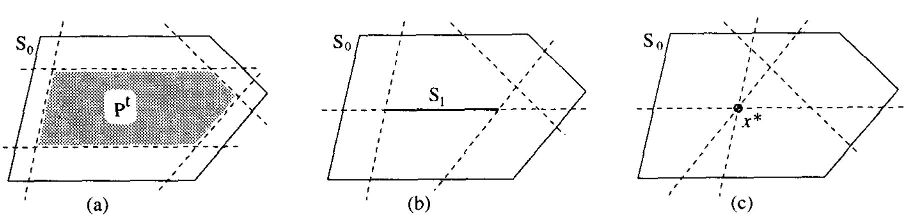


Fig. 2. The centroid of $S_0$ is obtained by "pushing in" all nonsaturated constraints (a) until some of them collapse (b). The process continues recursively down to $x^*$ (c).

**Remark.** The previous notion of center is of analytic nature, that is, it depends on the analytic description of $P$. For instance, the multiplication of an inequality by a positive constant may change the centroid. Let us also notice that the polytope $S_j$ may collapse to a singleton for $j<k$. It suffices that $\{A_i\colon i\in I_j\}$ contain $n$ linearly independent vectors.

Let us now prove the announced convergence result.

**Proposition 3.2.** *The curve $x(r)$ converges towards the centroid $x^*$ of the optimal face $S_0$ as $r$ tends to zero*:

$$\lim_{r\downarrow 0} x(r)=x^* .$$

**Proof.** Since $x(r)$ stays bounded as $r$ goes to zero, it suffices to prove that $x^*$ is the unique accumulation point. Let $\tilde{x}:=\lim x_k$ where $x_k:=x(r_k)$ with $r_k\downarrow 0$ be an accumulation point, and define $x_k^*:=x_k+(x^*-\tilde{x})$ so that $x_k^*\to x^*$. The optimality of $x_k$ gives

$$c'x_k+r_k\sum_{i\in I} \mathrm{e}^{(A_ix_k-b_i)/r_k}\leqslant c'x_k^*+r_k\sum_{i\in I} \mathrm{e}^{(A_ix_k^*-b_i)/r_k}$$

but since $\tilde{x}$, $x^*\in S_0$ we have $c'x_k=c'x_k^*$ and $A_ix_k=A_ix_k^*$ for all $i\in I_0$, and we get

$$\sum_{i\notin I_0} \exp[(A_ix_k-b_i)/r_k]\leqslant \sum_{i\notin I_0} \exp[(A_ix_k^*-b_i)/r_k] . \tag{5}$$

It follows that

$$\exp[-\phi_0(x_k)/r_k]\leqslant m\exp[-\phi_0(x_k^*)/r_k] ,$$

hence $\phi_0(x_k)\geqslant\phi_0(x_k^*)-r_k\ln m$, and taking limits we deduce $\phi_0(\tilde{x})\geqslant\phi_0(x^*)=t_1$, showing that $\tilde{x}\in S_1$.

We proceed by noting that $\tilde{x}\in S_1$ implies $A_i\tilde{x}=A_ix^*$ for $i\in J_1$, hence $A_ix_k=A_ix_k^*$ for these $i$'s. Eliminating the corresponding terms from (5) we get

$$\sum_{i\notin I_1} \exp[(A_ix_k-b_i)/r_k]\leqslant \sum_{i\notin I_1} \exp[(A_ix_k^*-b_i)/r_k] ,$$

and similarly as above we get $\phi_1(x_k)\geqslant\phi_1(x_k^*)-r_k\ln m$ from which $\tilde{x}\in S_2$, and so on.

By induction we ultimately get $\tilde{x}\in S_k=\{x^*\}$, proving the result. □

## 4. Convergence of the dual trajectory

We prove next that the dual trajectory $\lambda(r)$ converges as $r\downarrow 0$ towards an optimal solution of (D). More precisely, let $\lambda^*$ be the unique solution of the strictly convex problem

$$(\mathrm{D}_0) \qquad \min\left\{\sum_{i\in I_0} \lambda_i(\ln\lambda_i - 1) : \lambda\in D^*\right\},$$

where $D^*$ denotes the optimal solution set of the dual problem (D), that is,

$$D^* = \{\lambda : A'\lambda = -c;\ \lambda_i = 0,\ i\notin I_0;\ \lambda_i \geqslant 0,\ i\in I_0\}.$$

**Proposition 4.1.** *The multipliers $\lambda(r)$ converge towards $\lambda^*$ as r tends to zero*:

$$\lim_{r\downarrow 0} \lambda(r) = \lambda^*.$$

**Proof.** By optimality of $\lambda(r)$ we have

$$b'\lambda(r) + r\sum_{i\in I} \lambda_i(r)(\ln\lambda_i(r) - 1) \leqslant b'\lambda^* + r\sum_{i\in I} \lambda_i^*(\ln\lambda_i^* - 1),$$

but since $\lambda(r)$ is feasible for (D) while $\lambda^*$ is optimal for the same problem, we have $b'\lambda(r) \geqslant b'\lambda^*$ and then

$$\sum_{i\in I} \lambda_i(r)(\ln\lambda_i(r) - 1) \leqslant \sum_{i\in I_0} \lambda_i^*(\ln\lambda_i^* - 1). \tag{6}$$

Thus the multipliers $\lambda_i(r)$ must stay bounded as $r$ goes to zero, since the function $t\to t(\ln t - 1)$ is coercive on $[0, +\infty)$.

Let us verify that $\lambda^*$ is the only accumulation point of $\lambda(r)$. Indeed, let $\tilde{\lambda}$ be an accumulation point and take $r_k\downarrow 0$ so that $\lambda(r_k)\to\tilde{\lambda}$. Since $\tilde{\lambda}_i = \lim_k \exp[(A_i x(r_k) - b_i)/r_k] = 0$ for $i\notin I_0$ we obtain $\tilde{\lambda}\in D^*$, and passing to the limit in (6) we deduce $\tilde{\lambda}$ is an optimal solution of $(\mathrm{D}_0)$, hence $\tilde{\lambda} = \lambda^*$ as claimed. □

The next result proves that $\lambda^*$ is in fact a *central* point of the dual optimal face.

**Proposition 4.2.** *For all $i\in I_0$ one has $\lambda_i^* > 0$. More precisely, there exists $d\in\mathbb{R}^n$ such that $\lambda_i^* = \exp(A_i d)$ for $i\in I_0$.*

**Proof.** Suppose $\lambda_i^* = 0$ for some $i\in I_0$ and choose a dual optimal solution with $\lambda_i > 0$ (see e.g. [18, p. 95]). Then, the directional derivative of $\theta(\lambda) := \sum_{j\in I_0}\lambda_j(\ln\lambda_j - 1)$ at $\lambda^*$ in the direction $(\tilde{\lambda} - \lambda^*)$ is $-\infty$, contradicting the fact that $\lambda^*$ minimizes $\theta$ among all dual optimal solutions.

Since $\ln\lambda_i(r) = A_i(x(r) - x^*)/r$ for all $i\in I_0$, the vector $(\ln\lambda_i(r) : i\in I_0)$ belongs to the range of the linear operator $(A_i\colon i\in I_0)$. This property is inherited by the limit vector $(\ln\lambda_i^* : i\in I_0)$ and the result follows. □

## 5. Asymptotic analysis of trajectories near zero

Proposition 4.1 proves the convergence of the multipliers. As a matter of fact, these multipliers converge "very fast", in the sense made precise by the following definition.

**Definition 2.** We say that $\eta: [0, +\infty) \to \mathbb{R}^k$ converges *exponentially* towards zero if there exists $\mu > 0$ such that

$$\lim_{r\downarrow 0} \frac{\|\eta(r)\|}{\exp(-\mu/r)} = 0 .$$

We shall denote this fact by $\eta(r) \sim e(r)$.

Notice that if $\eta(r) \sim e(r)$, then we also have $\eta(r)/r^q \sim e(r)$ for all $q \geqslant 0$.

**Proposition 5.1.** *We have* $\dot{\lambda}(r) \sim e(r)$ *so that*

$$\lambda(r) = \lambda^* + \nu(r)$$

*with* $\nu(r)$ *converging exponentially to* 0. *In particular,*

$$\dot{\lambda}(0) := \lim_{r\downarrow 0} \frac{\lambda(r) - \lambda^*}{r} = \lim_{r\downarrow 0} \dot{\lambda}(r) = 0 .$$

**Proof.** From Eq. (1) we have $A'DA\dot{x}(r) = A'D\ln\lambda$, so that $\dot{x}(r)$ solves the weighted least-squares problem

$$\min_z (Az - \ln\lambda)'D(Az - \ln\lambda)$$

that is,

$$\min_z \sum_{i\in I} \lambda_i(r)[A_i z - \ln\lambda_i(r)]^2 . \tag{7}$$

In particular, taking $z := (x(r) - x^*)/r$ we obtain

$$\lambda_i(r)[A_i\dot{x}(r) - \ln\lambda_i(r)]^2 \leqslant \frac{1}{r^2} \sum_{j\notin I_0} \lambda_j(r)[A_j x^* - b_j]^2 \leqslant \frac{K}{r^2} \max_{j\notin I_0} \lambda_j(r) ,$$

where $K = \sum_{j\notin I_0}[A_j x^* - b_j]^2$.

We observe next that $(\mathrm{d}/\mathrm{d}r)[r\ln\lambda_i(r)] = A_i\dot{x}(r)$ so that

$$r\frac{\dot{\lambda}_i(r)}{\lambda_i(r)} = A_i\dot{x}(r) - \ln\lambda_i(r) ,$$

and therefore

$$\dot{\lambda}_i^2(r) = \frac{1}{r^2}\lambda_i^2(r)[A_i\dot{x}(r) - \ln\lambda_i(r)]^2 \leqslant \frac{K}{r^4}\lambda_i(r)\max_{j\notin I_0}\lambda_j(r) .$$

It follows that for some constant $\tilde{K}$ and $r$ small

$$|\dot{\lambda}_i(r)| \leqslant \frac{\tilde{K}}{r^2}\sqrt{\max_{j\notin I_0}\lambda_j(r)} ,$$

and since $\lambda_j(r) \leqslant \exp(-\tilde{t}/r)$ for any $\tilde{t}<t_1, j\notin I_0$ and $r$ small, we may take any $\mu < \frac{1}{2}t_1$ to check the exponential convergence of $\dot{\lambda}(r)$ towards zero. The rest of the proposition follows easily from this. $\square$

Our next step is to prove the existence of a right derivative of $x(r)$ at $r=0$,

$$d^* = \lim_{r\downarrow 0}\frac{x(r) - x^*}{r} .$$

To this end we shall exploit the fact that, from the very definition of $x(r)$, the vector $d(r) := (x(r) - x^*)/r$ is the unique minimizer of the function

$$\Theta_r(d) = c'd + \sum_{i\in I_0} e^{A_i d} + e^{-t_1/r}\sum_{i\in J_1} e^{A_i d} + \cdots + e^{-t_k/r}\sum_{i\in J_k} e^{A_i d} .$$

The next simple linear algebra result will be useful in what follows.

**Lemma 5.2.** *Assume $h(r)\in\mathbb{R}^n$ is such that $A_ih(r)$ is bounded (respectively $A_ih(r)\sim e(r)$) for all $i\in I_j$. Then, the projection $h_j(r)$ of $h(r)$ onto the space $E_j := \operatorname{span}\{A_i' \colon i\in I_j\}$ is bounded (respectively $h_j(r)\sim e(r)$).*

**Proof.** Choose a basis for $E_j$ out of $\{A_i' \colon i\in I_j\}$ and form a matrix $B$. Then $h_j(r) = B(B'B)^{-1}B'h(r)$ from which the result follows. $\square$

**Proposition 5.3.** *The vector $d(r)$ stays bounded as $r$ tends to zero.*

**Proof.** Let us prove by induction that the projection $d_j(r)$ of $d(r)$ onto the space $E_j := \operatorname{span}\{A_i' \colon i\in I_j\}$ is bounded. Since $E_k = \mathbb{R}^n$ the conclusion will follow.

Since $A_i d(r) \to \ln\lambda_i^*$ for all $i\in I_0$, Lemma 5.2 implies that $d_0(r)$ is bounded. Moreover, since $c = -\sum_{i\in I_0}\lambda_i^* A_i' \in E_0$, we also get that $c'd(r)$ is bounded.

Suppose that $d_{j-1}(r)$ is bounded and let us show the same holds for $d_j(r)$.

We begin by proving that $A_i d(r)$ is bounded above for all $i\in I_j$. The induction hypothesis shows that $A_i d(r) = A_i d_{j-1}(r)$ is bounded for all $i\in I_{j-1}$. Moreover, since $c\in E_0\subset E_{j-1}$, we have $c'd(r) = c'd_{j-1}(r)$ so we may cancel the first terms in the inequality $\Theta_r(d(r)) \leqslant \Theta_r(d_{j-1}(r))$ in order to deduce

$$\sum_{J_j} e^{A_i d(r)} \leqslant \sum_{J_j} e^{A_i d_{j-1}(r)} + e^{(t_j - t_{j+1})/r}\sum_{J_{j+1}} e^{A_i d_{j-1}(r)} + \cdots + e^{(t_j - t_k)/r}\sum_{J_k} e^{A_i d_{j-1}(r)} .$$

From (3) and the induction hypothesis the right hand side above stays bounded, so we deduce that $A_i d(r)$ is also bounded above for all $i \in J_j = I_j \setminus I_{j-1}$.

Now, if $d_j(r)$ were not bounded we could find $r_k \downarrow 0$ such that $\|d_j(r_k)\| \to \infty$. Then, since the quantities $A_i d_j(r) = A_i d(r)$ for all $i \in I_j$ and $c' d_j(r) = c' d(r)$ are bounded above, every accumulation point $e$ of $d_j(r_k) / \|d_j(r_k)\|$ belongs to $E_j$ and satisfies $A_i e \leqslant 0$, $i \in I_j$ and $c'e \leqslant 0$.

But then, from (4) and since $A_i x^* < b_i - t_j$ for $i \notin I_j$, we deduce that $x^* + te \in S_j$ for small enough $t > 0$. By the definition of $I_0, J_1, \ldots, J_j$ we deduce $A_i e = 0$ for all $i \in I_j$ and, since $e \in E_j$, this implies $e = 0$ in contradiction with $\|e\| = 1$. □

We may now identify the unique accumulation point of $d(r)$. To this end we shall use the following simple fact whose proof is left to the reader.

**Lemma 5.4.** *Let $\phi : \mathbb{R}^d \to \mathbb{R}$ be a strictly convex function, $H$ an affine subspace parallel to the linear subspace $E$, and $B$ a $d \times n$ matrix. If the problem*

$$\min_{z \in H} \phi(Bz)$$

*has a nonempty optimal solution set $L$, then for every $z_0 \in L$ we have*

$$L = z_0 + [E \cap \operatorname{Ker} B] .$$

Let us take $d^*$ an accumulation point of $d(r)$. Then we have $A_i d^* = \ln \lambda_i^*$ for all $i \in I_0$, and therefore

$$c + \sum_{i \in I_0} e^{A_i d^*} A_i' = 0 .$$

This proves that $d^*$ is an optimal solution of

$$(\mathrm{Q}_0) \qquad \min_d \, c'd + \sum_{i \in I_0} e^{A_i d}$$

whose optimal solution set is, by the previous lemma, $L_0 = d^* + \bigcap_{i \in I_0} \operatorname{Ker} A_i$.

Now, if we take any other $d_1 \in L_0$ and we consider $d_1(r) := d(r) + (d_1 - d^*)$, we have $A_i d_1(r) = A_i d(r)$ for all $i \in I_0$, and also $c' d_1(r) = c' d(r)$. Thus, canceling the first two terms in the inequality $\Theta_r(d(r)) \leqslant \Theta_r(d_1(r))$ we obtain

$$\sum_{i \in J_1} e^{A_i d(r)} \leqslant \sum_{i \in J_1} e^{A_i d_1(r)} + e^{(t_1 - t_2)/r} \sum_{i \in J_2} e^{A_i d_1(r)} + \cdots + e^{(t_1 - t_k)/r} \sum_{i \in J_k} e^{A_i d_1(r)}$$

and passing to the limit (in the subsequence $r_k$ such that $d(r_k) \to d^*$), we get

$$\sum_{i \in J_1} e^{A_i d^*} \leqslant \sum_{i \in J_1} e^{A_i d_1} .$$

This proves that $d^*$ is also an optimal solution of the problem

$$(\mathrm{Q}_1) \qquad \min_{d \in L_0} \sum_{i \in J_1} \mathrm{e}^{A_i d}$$

whose optimal solution set is $L_1 = d^* + \bigcap_{i \in I_1} \operatorname{Ker} A_i$.

We may proceed inductively by showing that $d^*$ is an optimal solution of

$$(\mathrm{Q}_2) \qquad \min_{d \in L_1} \sum_{i \in J_2} \mathrm{e}^{A_i d}$$

with optimal solution set $L_2 = d^* + \bigcap_{i \in I_2} \operatorname{Ker} A_i$, and so on.

Finally we shall get that $d^*$ solves

$$(\mathrm{Q}_k) \qquad \min_{d \in L_{k-1}} \sum_{i \in J_k} \mathrm{e}^{A_i d}$$

with optimal solution set

$$L_k = d^* + \bigcap_{i \in I_k} \operatorname{Ker} A_i = d^* + \bigcap_{i \in I} \operatorname{Ker} A_i = \{d^*\}\ .$$

**Proposition 5.5.** *Let $d^*$ be the unique solution of* $(\mathrm{Q}_k)$ *as above. Then*

$$\dot{x}(0) := \lim_{r \downarrow 0} \frac{x(r) - x^*}{r} = d^*\ .$$

**Proof.** It suffices to observe that every accumulation point of $d(r)$ belongs to $L_k$ which is reduced to a singleton $\{d^*\}$ by the full rank condition on $A$. □

To conclude this section we shall prove that the curve $x(r)$ is also $C^1$ at $r=0$, in the sense that $\dot{x}(r)$ tends to $\dot{x}(0)$ as $r$ decreases to 0.

**Proposition 5.6.** $\dot{d}(r) \sim e(r)$.

**Proof.** Since $\dot{d}(r) = (\dot{x}(r) - d(r))/r$, it suffices to prove that the vector $h(r) := r(d(r) - \dot{x}(r))$ converges to zero exponentially fast. To this end we notice that $\dot{x}(r)$ is the solution of (7), so that $h(r)$ minimizes the function

$$\Gamma_r(h) := \sum_{i \in I} \lambda_i(r)(A_i h + s_i)^2$$

where $s_i := A_i x^* - b_i$.

First of all we prove that $h(r)$ stays bounded as $r$ goes to zero, which by Lemma 5.2 amounts to showing that $A_i h(r)$ stays bounded for all $i \in I$. Let us do this by induction. From the inequality $\Gamma_r(h(r)) \leqslant \Gamma_r(0)$ we deduce that $A_i h(r) \sim e(r)$ for all $i \in I_0$, hence these $A_i h(r)$ are bounded. Assume now that $A_i h(r)$ stays bounded for all $i \in I_j$, and consider

$h_j(r)$ the projection of $h(r)$ onto $E_j$, so that $h_j(r)$ is bounded by Lemma 5.2. The inequality $\Gamma_r(h(r)) \leqslant \Gamma_r(h_j(r))$ implies

$$\sum_{i \notin I_j} \lambda_i(r)(A_i h(r) + s_i)^2 \leqslant \sum_{i \notin I_j} \lambda_i(r)(A_i h_j(r) + s_i)^2 .$$

We notice that $\lambda_i(r) = \exp[A_i d(r)] \exp(-t_{j+1}/r)$ for $i \in J_{j+1}$. Thus, multiplying the previous inequality by $\exp(t_{j+1}/r)$ we observe that the right hand side will stay bounded by some constant $K$ so that for all $i \in J_{j+1}$ we obtain

$$(A_i h(r) + s_i)^2 \leqslant K \mathrm{e}^{-A_i d(r)} .$$

It follows that $A_i h(r)$ is also bounded for $i \in J_{j+1}$. This achieves the induction step, and the boundedness of $h(r)$ has been established.

The second step is to prove that in fact $h(r) \sim e(r)$, which is again equivalent to showing that $A_i h(r) \sim e(r)$ for all $i \in I$. We already observed that this is the case for $i \in I_0$. Suppose the property holds for $i \in I_j$ so that, by Lemma 5.2, $h_j(r) \sim e(r)$. Then, by an argument similar to the one used to prove the boundedness of $h(r)$, and since $s_i = -t_{j+1}$ for all $i \in J_{j+1}$, we obtain

$$\sum_{i \in J_{j+1}} \mathrm{e}^{A_i d(r)} [A_i h(r)]^2 \leqslant \sum_{i \in J_{j+1}} \mathrm{e}^{A_i d(r)} \{ [A_i h_j(r)]^2 - 2t_{j+1} A_i (h_j(r) - h(r)) \} + E(r) \tag{8}$$

where $E(r) \sim e(r)$. Moreover,

$$c + \sum_{i \in I} \mathrm{e}^{A_i d(r)} \mathrm{e}^{s_i/r} A_i' = 0$$

and since $(h_j(r) - h(r))$ is bounded and orthogonal to all the $A_i$'s for $i \in I_j$ (and thus also orthogonal to $c$), we obtain

$$\sum_{i \in J_{j+1}} \mathrm{e}^{A_i d(r)} A_i (h_j(r) - h(r)) = -\mathrm{e}^{t_{j+1}/r} \sum_{i \notin I_{j+1}} \mathrm{e}^{A_i d(r)} \mathrm{e}^{s_i/r} A_i (h_j(r) - h(r)) \sim e(r)$$

as $s_i + t_{j+1} < 0$ for $i \notin I_{j+1}$. Therefore, the right hand side of (8) decays exponentially and we deduce $A_i h(r) \sim e(r)$ for all $i \in J_{j+1}$. This completes the induction step and we have proved $h(r) \sim e(r)$ as required. □

**Proposition 5.7.** *The following vectors converge exponentially towards zero,*

(a) $(d(r) - d^*) \sim e(r)$.

(b) $(\dot{x}(r) - d^*) \sim e(r)$.

(c) $(x(r) - x^* - rd^*) \sim e(r)$.

*In particular* $x(r)$ *is a* $C^1$ *curve for* $r \geqslant 0$.

**Proof.** Assertion (a) follows from the mean value theorem. Part (b) is a consequence of the equality $\dot{x}(r) = r\dot{d}(r) + d(r)$, while part (c) follows from (a). The last assertion is obvious from (b). □

We summarize the main results of this section in the next theorem.

**Theorem 5.8.** *The primal and dual trajectories have asymptotic expansions*

$$x(r) = x^* + rd^* + \eta(r) ,$$
$$\lambda(r) = \lambda^* + \nu(r) ,$$

*where the error terms $\eta(r)$ and $\nu(r)$ converge exponentially fast to* 0 *as r tends to* 0.

## 6. Higher order differentiability

So far we have shown that the primal and dual trajectories are of class $C^\infty$ on $(0, +\infty)$, and also right differentiable at $r=0$ with the derivatives being continuous at this point. Concerning higher order derivatives at $r=0$, the results in the previous section imply

$$\ddot{x}(0) := \lim_{r\downarrow 0} \frac{\dot{x}(r) - \dot{x}(0)}{r} = 0 ,$$
$$\ddot{\lambda}(0) := \lim_{r\downarrow 0} \frac{\dot{\lambda}(r) - \dot{\lambda}(0)}{r} = 0 .$$

More generally, all higher order Taylor expansions have only linear terms

$$x(r) = x(0) + r\dot{x}(0) + o(r^q) ,$$
$$\lambda(r) = \lambda(0) + o(r^q) .$$

However, we may not assert that the trajectory is of class $C^\infty$ at $r=0$, nor even $C^2$ since we have not proved that $\lim \ddot{x}(r) = 0$. While such a property is plausible, a direct attempt to prove it in a similar way to what was done for $\dot{x}(r)$ would be very intrincate. We shall adopt a different strategy which allows us to prove the trajectories are $C^\infty$ at $r=0$ (and even for $r<0$), but under the rather restrictive hypothesis

(H) (LP) and its dual (D) have unique solutions ,

which amounts to $B := \{A_i' \colon i \in I_0\}$ being a basis of $\mathbb{R}^n$.

To this end we rewrite the equation

$$c + \sum_{i\in I} \exp[(A_i x - b_i)/r] A_i' = 0 \tag{9}$$

which characterizes $x(r)$, as the equivalent system

$$\text{(S)} \qquad \begin{cases} c + \Sigma \lambda_i A_i' = 0 , & \\ A_i x - b_i - r \ln \lambda_i = 0 & \text{for } i \in I_0 , \\ \beta(A_i x - b_i, r) - \lambda_i = 0 & \text{for } i \notin I_0 , \end{cases}$$

where $\beta: ]-\infty, 0[\times\mathbb{R}\to\mathbb{R}$ is the $C^\infty$ function defined by

$$\beta(s, r)=\begin{cases}\exp(s/r) & \text{if } r>0,\\ 0 & \text{if } r\leqslant 0.\end{cases}$$

System (S) is equivalent to (9) for $r>0$ but is also meaningful for $r=0$ (and even for $r<0$). The Jacobian of this system with respect to the pair $(x, \lambda)$ at the point $(x^*, \lambda^*, 0)$ is

$$J^*=\begin{pmatrix}0 & B & (A_i')_{i\notin I_0}\\ B' & 0 & 0\\ 0 & 0 & -\mathrm{Id}\end{pmatrix}$$

where Id denotes the identity matrix of dimension $|I\setminus I_0|$. $J^*$ is clearly nonsingular under hypothesis (H), so we can apply the Implicit Function Theorem to deduce

**Proposition 6.1.** *Assuming* (H), *the primal and dual trajectories* $x(r)$ *and* $\lambda(r)$ *are of class* $C^\infty$ *on* $[0, +\infty)$, *with*

$$x^{(k)}(0)=0 \quad \textit{for } \mathrm{k}\geqslant 2,$$
$$\lambda^{(k)}(0)=0 \quad \textit{for } \mathrm{k}\geqslant 1.$$

**Proof.** Immediate by the previous discussion. □

This result provides further support to the conjecture that the trajectories are $C^\infty$ at $r=0$, but the proof in the case of multiple primal or dual solutions remains an open question.

## 7. Asymptotic behavior of trajectories at infinity

We supplement the previous results by studying the behavior of the primal and dual trajectories when $r$ tends to $\infty$. To this end we consider the unconstrained problem

$$(\mathrm{P}_\infty) \qquad \min_d c'd+\sum_{i\in I}\exp(A_i d)$$

and the perturbation function $\varphi(d, u):=c'd+\sum\exp(A_i d+u_i)$ which gives the dual

$$(\mathrm{D}_\infty) \qquad \min_\lambda\left(\sum_{i\in I}\lambda_i(\ln\lambda_i-1): A'\lambda+c=0,\ \lambda\geqslant 0\right).$$

**Proposition 7.1.** *Problems* $(\mathrm{P}_\infty)$ *and* $(\mathrm{D}_\infty)$ *are strictly convex and coercive, their optimal values satisfy* $v(\mathrm{P}_\infty)+v(\mathrm{D}_\infty)=0$, *and the corresponding unique solutions* $d^\infty$ *and* $\lambda^\infty$ *are related by*

$$\lambda_i^\infty=\exp(A_i d^\infty).$$

**Proof.** It suffices to apply Proposition 2.1 with $b=0$ and $r=1$. □

**Proposition 7.2.** *With the previous notation we have*

$$\lim_{r\uparrow\infty} \frac{x(r)}{r} = \lim_{r\uparrow\infty} \dot{x}(r) = d^\infty , \tag{10}$$

*and also*

$$\lim_{r\uparrow\infty} \lambda(r) = \lambda^\infty . \tag{11}$$

**Proof.** Let us define $\hat{d}(r) := x(r)/r$. The inequality $f_r(x(r)) \leqslant f_r(rd^\infty)$ gives

$$c'\hat{d}(r) + \sum_{i\in I} \exp(A_i\hat{d}(r))\exp(-b_i/r) \leqslant c'd^\infty + \sum_{i\in I} \exp(A_i d^\infty)\exp(-b_i/r) , \tag{12}$$

and since for $r$ sufficiently large $\frac{1}{2} \leqslant \exp(-b_i/r) \leqslant 2$, we deduce

$$c'\hat{d}(r) + \tfrac{1}{2} \sum_{i\in I} \exp(A_i\hat{d}(r)) \leqslant c'd^\infty + 2 \sum_{i\in I} \exp(A_i d^\infty) .$$

Using the previous proposition (with $2c$ instead of $c$) we observe that the function on the left has bounded level sets, so that $\hat{d}(r)$ must be bounded. Then, passing to the limit in (12) we deduce that every accumulation point of $\hat{d}(r)$ solves ($\mathrm{P}_\infty$), hence $\lim \hat{d}(r) = d^\infty$ proving the first half of (10).

Property (11) follows immediately from this since $\lambda_i(r) = \exp[A_i\hat{d}(r) - b_i/r]$. Moreover, (1) implies

$$\dot{x}(r) = (A'DA)^{-1}A'D \frac{Ax(r)-b}{r} = \hat{d}(r) - \frac{1}{r}(A'DA)^{-1}A'Db$$

and then we also have $\lim \dot{x}(r) = d^\infty$, completing the proof of (10). □

The previous result shows that the primal trajectory diverges when $r\uparrow\infty$ (except when $d^\infty = 0$). We conclude by showing that $x(r)$ admits an asymptotic ray at infinity, and by studying the asymptotic expansion of $\lambda(r)$.

To this end let $D^\infty := \mathrm{diag}\{\lambda_i^\infty : i\in I\}$ and consider $x^\infty = (A'D^\infty A)^{-1}A'D^\infty b$, the unique solution of the following weighted least-squares problem

$$\text{(WLS)} \qquad \min_x \ (Ax-b)'D^\infty(Ax-b) .$$

**Theorem 7.3.** *The primal trajectory has the asymptotic behavior*

$$x(r) = x^\infty + rd^\infty + \rho(r) \tag{13}$$

*with* $\lim_{r\uparrow\infty}\rho(r)=0$, *and the dual trajectory has an expansion of the form*

$$\lambda_i(r)=\lambda_i^\infty[1+\delta_i^\infty/r+\xi(r)] \tag{14}$$

*with* $\delta_i^\infty:=A_ix^\infty-b_i$ *and* $\lim_{r\uparrow\infty}r\xi(r)=0$.

**Proof.** Let $z(r):=x(r)-rd^\infty$. The optimality of $x(r)$ for $f_r$ proves that $z(r)$ is the unique solution of

$$\min_z c'z+r\sum_{i\in I}\lambda_i^\infty\exp[(A_iz-b_i)/r]\ .$$

Noting that $c'=-\Sigma\lambda_i^\infty A_i$, adding the constant $r\Sigma\lambda_i^\infty(b_i/r-1)$ to the objective function and multiplying it by $r$, we deduce that $z(r)$ also minimizes the function

$$\begin{aligned}\Omega_r(z)&:=r^2\sum_{i\in I}\lambda_i^\infty(\exp[(A_iz-b_i)/r]-1-[(A_iz-b_i)/r])\\&=\sum_{i\in I}\lambda_i^\infty(A_iz-b_i)^2\omega((A_iz-b_i)/r)\ ,\end{aligned}$$

where $\omega(t):=\Sigma_{k=0}^\infty t^k/(k+2)!$. Since $z(r)/r$ tends to 0, for each $\epsilon>0$ we have $\omega((A_iz(r)-b_i)/r)\geqslant\omega(0)-\epsilon=\frac{1}{2}-\epsilon$ for all $r$ large enough. Thus, using the inequality $\Omega_r(z(r))\leqslant\Omega_r(x^\infty)$, we find

$$(\tfrac{1}{2}-\epsilon)\sum_{i\in I}\lambda_i^\infty(A_iz(r)-b_i)^2\leqslant\sum_{i\in I}\lambda_i^\infty(A_ix^\infty-b_i)^2\omega((A_ix^\infty-b_i)/r)\ .$$

From this we conclude that $Az(r)$, and a fortiori $z(r)$, stays bounded as $r$ tends to $\infty$. Moreover, passing to the limit in the last inequality, we deduce that every accumulation point $\bar z$ of $z(r)$ satisfies,

$$(\tfrac{1}{2}-\epsilon)\sum_{i\in I}\lambda_i^\infty(A_i\bar z-b_i)^2\leqslant\tfrac{1}{2}\sum_{i\in I}\lambda_i^\infty(A_ix^\infty-b_i)^2\ .$$

Letting $\epsilon$ tend to 0 we deduce that $\bar z$ solves (WLS), hence $\bar z=x^\infty$ proving (13).

The expansion (14) follows from (13) since

$$\lambda_i(r)=\exp[(A_ix(r)-b_i)/r]=\lambda_i^\infty\exp[(\delta_i^\infty+A_i\rho(r))/r]\ . \qquad\square$$

**Remark.** Notice that when $d^\infty=0$, the trajectory $x(r)$ does not diverge but converges to $x^\infty$ when $r\uparrow\infty$. This happens if and only if the vector $e=(1,\ldots,1)'$ is a dual feasible solution. As a matter of fact, if we know any strictly positive dual feasible solution we may scale the dual variables so that $e$ becomes feasible, and then we will have $d^\infty=0$, $\lambda^\infty=e$, and $x^\infty=(A'A)^{-1}A'b$ the unique solution of

$$\min_x \|Ax-b\|^2\ .$$

## References

[1] D. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods* (Academic Press, New York, 1982).

[2] J. Borwein and A. Lewis, "A survey of convergence results for maximum entropy methods," CORR 92-32, University of Waterloo (1992).

[3] Y. Censor et al., "On iterative methods for linearly constrained entropy maximization," *Numerical Analysis and Mathematical Modelling*, Banach Center Publ. 24 (1990) pp. 145–163.

[4] R. Cominetti and J.P. Dussault, "A stable exponential-penalty method with superlinear convergence," to appear in *Journal of Optimization Theory and Applications.*

[5] D. den Hertog, C. Roos and T. Terlaky, "Inverse barrier methods for linear programming," Report 91-27, Faculty of Technical Mathematics and Informatics, Delft University of Technology, The Netherlands (1991).

[6] T. Elfving, "On some methods for entropy maximization and matrix scaling," *Linear Algebra and Its Applications* 34 (1980) 321–339.

[7] S. Erlander, "Entropy in linear programs," *Mathematical Programming* 21 (1981) 137–151.

[8] K.R. Frisch, "The logarithmic potential method of convex programming," Memorandum of May 13, 1955, University Institute of Economics, Oslo.

[9] C. Gonzaga, "Path following methods for linear programming," *SIAM Review* 34(2) (1992) 167–224.

[10] O. Güler, "Limiting behavior of weighted central paths in linear programming," Technical Report, Department of Mathematics and Statistics, University of Maryland, Baltimore, MD 21228 (1991).

[11] F. Jarre, "Interior point methods for convex programming," *Applied Mathematics and Optimization* 26 (1992) 287–311.

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

[13] L.G. Khachiyan, "A polynomial algorithm for linear programming," *Soviet Mathematics Doklady* 20 (1979) 191–194.

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

[15] K. Mouallif and P. Tossings, "Une méthode de pénalisation exponentielle associée à une régularisation proximale," *Bulletin de la Société Royale des Sciences de Liège* 56 (1987) 181–190.

[16] Yu. Nesterov and A. Nemirovsky, *Self-Concordant Function and Polynomial Time Methods in Convex Programming*, Moscow, USSR (1989).

[17] R.T. Rockafellar, *Conjugate Duality and Optimization*, CBMS-NFS Regional Conference Series in Applied Mathematics (SIAM, Philadelphia, 1974).

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

[19] J.J. Strodiot and V.H. Nguyen, "An exponential penalty method for non-differentiable minimax problems with general constraints," *JOTA* 27 (1979) 205–219.