NEW PROOFS OF CONVERGENCE FOR THE DUAL AFFINE SCALING ALGORITHM

Walter F. Mascarenhas

Novembro RP 69/93

Relatório de Pesquisa

Instituto de Matemática Estatística e Ciência da Computação

UNIVERSIDADE ESTADUAL DE CAMPINAS
Campinas - São Paulo - Brasil

ABSTRACT – We present two ideas that help in the analysis of the dual affine scaling algorithm: sorting the slacks and taking a QR factorization of the constrains. Using these ideas, we prove that the iterates always converge. The proof holds if at each iteration we move an arbitrary fraction of the step to the boundary of the feasible region and it needs no hypothesis on degeneracy. However, it does not show that the iterates converge to an optimal solution. We present a new proof of convergence to the optimum for primal nondegenerate programs and for programs with two variables.

IMECC – UNICAMP
Universidade Estadual de Campinas
CP 6065
13081-970 Campinas SP
Brasil

Novembro – 1993

New proofs of convergence for the dual affine scaling algorithm

Walter F. Mascarenhas
Universidade Estadual de Campinas
Campinas S.P., Brazil

November 25, 1993

Abstract

We present two ideas that help in the analysis of the dual affine scaling algorithm: sorting the slacks and taking a QR factorization of the constrains. Using these ideas, we prove that the iterates always converge. The proof holds if at each iteration we move an arbitrary fraction of the step to the boundary of the feasible region and it needs no hypothesis on degeneracy. However, it does not show that the iterates converge to an optimal solution. We present a new proof of convergence to the optimum for primal nondegenerate programs and for programs with two variables.

key words. affine scaling algorithm, convergence, degeneracy.

1. Introduction

We discuss the convergence of the dual affine scaling algorithm [T], used for solving linear programs $\Pi(A, b, c)$ of the form

$$\text{minimize} \quad z(x) = c^t x, \qquad \text{subject to} \quad A^t x \geq b. \tag{1.1}$$

The matrix A is $n \times m$ ($n \leq m$.) The feasible region of Π is $\mathcal{F}_\Pi = \{ x \in I\!R^n \text{ s.t. } A^t x \geq b \}$ and $\mathcal{F}_\Pi^+$ is its interior. As [MTW], we are interested in programs Π with an optimal solution and such that $c \neq 0$, $\mathcal{F}_\Pi^+ \neq \emptyset$, and A has full rank. We call such programs *acceptable*. We denote the fraction of the step to the boundary taken at each iteration by λ. We assume that the reader is familiar with the dual affine scaling algorithm. In resume, the results of [MTW], when adapted to the dual affine scaling algorithm, show that the iterates converge regardless of degeneracy of Π and the size of λ. With the additional assumption that $\lambda \leq 2/3$, it is also shown that the dual estimates converge and the iterates converge to an optimal solution. In [M] we present a primal degenerate example in which some iterates converge to a vertex that is not optimal if $\lambda = 0.999$.

In this paper, we present a new proof of convergence for arbitrary programs and λ's and prove convergence to the optimum for primal nondegenerate programs and for programs with only two variables. We hope that our ideas can help in the analysis of the behavior of other interior point methods near the boundary of $\mathcal{F}_\Pi$, which we call $\partial\mathcal{F}_\Pi$. Things get complicated near $\partial\mathcal{F}_\Pi$ because the algorithm steps in the direction $d(\Pi, x) = (A[A^t x - b]^{-2}A^t)^{-1}c$, where $[v]$ is the diagonal matrix with diagonal equal to v. Some of the slacks $\xi = A^t x - b$ vanish at $\partial\mathcal{F}_\Pi$ and $d(\Pi, x)$ is singular there. This is the main difficulty in the analysis of the algorithm. To simplify this analysis, we introduce the permutation P such that $P\xi$ is sorted in increasing order and the QR factorization $AP^t = QR$. The direction $d(\Pi, x)$ can then be written as

$$d(\Pi, x) = Q(R[P\xi]^{-2}R^t)^{-1}Q^t c. \tag{1.2}$$

This equation shows that we do not loose generality if we assume that the slacks are sorted in increasing order and A is upper triangular, if we want to analyze a single step. (This argument does not work for multiple steps because the order of the slacks may change from one iteration to the next.) Using (1.2) we prove the following lemma:

Lemma 1. *If the program Π is acceptable, then there exists $K_\Pi \in I\!R$ such that $\|d(\Pi, x)\| \leq K_\Pi \; c^t d(\Pi, x)$ for all $x \in \mathcal{F}_\Pi^+$.*

Lemma 1 relates $\|x^{k+1} - x^k\|$ to the decay of the objective function, which cannot be too big because $c^t x^k$ is bounded from below. The reference [MTW] presents a different proof of Lemma 1 and shows how it implies the following Theorem:

Theorem 1. *For any acceptable program Π, $x^0 \in \mathcal{F}_\Pi^+$ and $\lambda \in (0,1)$, the sequence $\{x^k\}$ converges to $\bar{x}(x^0, \lambda, \Pi)$ in the boundary of $\mathcal{F}$.*

As it can be seen in [MTW], it is straightforward to use Lemma 1 to show that x^k is a Cauchy sequence and prove Theorem 1. Therefore, we will not present a proof of this Theorem. Our proof of Lemma 1, in section 2, is self-contained. Theorem 1 does not claim that $\bar{x}$ is optimum. We can guarantee the optimality of $\bar{x}$ for all $\lambda \in (0,1)$ for primal nondegenerate problems:

Theorem 2. *If the program Π is acceptable, $x^0 \in \mathcal{F}_\Pi^+$, $\lambda \in (0,1)$, and the bind-*

ing restrictions at $\bar{x}(x^0, \lambda, \Pi)$ are linearly independent, then $\bar{x}(x^0, \lambda, \Pi)$ is optimal.

Theorem 2 is proved in section 3. In section 4 we prove convergence for problems with only two variables:

Theorem 3. *If the program Π is acceptable, $x^0 \in \mathcal{F}_\Pi^+ \subset I\!R^2$ and $\lambda \in (0, 1)$, then $\bar{x}(x^0, \lambda, \Pi)$ is optimal.*

The example of convergence to a non-optimal solution in [M] has three variables and five restrictions. Therefore, this example is minimal and the three theorems above are the best results we can hope for the convergence of x^k for an arbitrary $\lambda \in (0, 1)$. The appendix of [T] shows how to adapt these results to the primal affine scaling algorithm.

2. Convergence

Proof of Lemma 1. In this proof, x and Π are fixed and we omit their dependencies from several functions. We start by taking a permutation $P \in I\!R^{m \times m}$ such that $P\xi$ is sorted in increasing order, that is, $\psi = P\xi$ satisfies $\psi_{i+1} \geq \psi_i$. Next, we take the QR factorization of AP^t, i.e., $A = QRP$, with Q orthogonal and R upper triangular. Notice that rank(R) = rank(A). Since Π is acceptable, R has full rank.

The statements in this proof are easy to verify in examples. However, the details needed to cover all cases are tedious. To motivate our arguments, we will illustrate them with the example

$$R = \begin{pmatrix} 0 & r_{12} & r_{13} & r_{14} & r_{15} & r_{16} \\ 0 & 0 & r_{23} & r_{24} & r_{25} & r_{26} \\ 0 & 0 & 0 & 0 & r_{25} & r_{36} \end{pmatrix}. \qquad (2.3)$$

The r_{ij} indicated above are different from 0. This example is quite general and if the reader understands the proof in this case, then it should not be hard to extended it to arbitrary R's. Although we use the example as motivation, the proof bellow is general.

For each $i \leq n$, the ith row of R has a first non-zero entry. We call $j(R, i)$ the index j of this entry. Formally, $r_{ij(R,i)} \neq 0$ and if $j < j(R, i)$ then $r_{ij} = 0$. Notice that $j(R, i+1) > j(R, i) \geq i$. Let $\phi \in I\!R^n$ be given by $\phi_i = \psi_{j(R,i)}$. From (1.2) we get

$$d = Q[\phi] \left([\phi]R[\psi]^{-2}R^t[\phi] \right)^{-1} [\phi]Q^tc = Q[\phi] \left(BB^t\right)^{-1} [\phi]Q^tc. \qquad (2.4)$$

where $B = [\phi]R[\psi]^{-1}$. In the example (2.3), $j(R, 1) = 2, j(R, 2) = 3$, and $j(R, 3) =$

5. The vector ϕ equals $(\psi_2, \psi_3, \psi_5)^t$ and

$$B = \begin{pmatrix} 0 & r_{12} & r_{13}\frac{\psi_2}{\psi_3} & r_{14}\frac{\psi_2}{\psi_4} & r_{15}\frac{\psi_2}{\psi_5} & r_{16}\frac{\psi_2}{\psi_6} \\ 0 & 0 & r_{23} & r_{24}\frac{\psi_3}{\psi_4} & r_{25}\frac{\psi_3}{\psi_5} & r_{26}\frac{\psi_3}{\psi_6} \\ 0 & 0 & 0 & 0 & r_{35} & r_{26}\frac{\psi_5}{\psi_6} \end{pmatrix}.$$

Notice that if $j < j(R,i)$, then $b_{ij} = r_{ij} = 0$. Otherwise,

$$b_{ij} = r_{ij}\frac{\phi_i}{\psi_j} = r_{ij}\frac{\psi_{j(R,i)}}{\psi_j}.$$

The vector ψ is sorted in increasing order and $j(R,i) \leq j$. Therefore, $b_{ij}^2 \leq r_{ij}^2$ and $\|B\|_F \leq \|R\|_F$. Here $\|M\|_F$ is the Frobenius norm of the matrix M, which is given by $\sqrt{\sum m_{ij}^2}$. Since Q and P are orthogonal, $\|R\|_F = \|A\|_F$ and the biggest singular value of B, $\sigma_{\max}(B)$, satisfies $\sigma_{\max}(B) \leq \|B\|_F \leq \|R\|_F = \|A\|_F$. Let $B = U\Sigma V^t$ be the singular value decomposition of B. It follows from (2.4) that

$$c^t d = c^t Q[\phi]\; U\; \Sigma^{-2}\; U^t\; [\phi]Q^t c \geq \frac{\|\; U[\phi]Q^t c\;\|^2}{\sigma_{\max}(B)^2} = \frac{\|\;[\phi]Q^t c\;\|^2}{\sigma_{\max}(B)^2} \geq \frac{\phi_\mu^2\,\eta^2}{\|A\|_F^2}, \tag{2.5}$$

where η is the smallest absolute value of a nonzero entry of $Q^t c$ and μ is the biggest index for which $(Q^t c)_\mu \neq 0$.

From (1.2) we get

$$d = \left(Q\;[\phi]^{-2}[\phi]^2\; RP[\xi]^{-2}P^t R^t Q^t \right)^{-1} c = Q \left([\phi]^2 R[\psi]^{-2} R^t \right)^{-1} [\phi]^2 Q^t c$$

and, calling $\sigma_{\min}(M)$ the smallest singular value of the matrix M,

$$\|d\| \leq \frac{\|\;[\phi]^2 Q^t c\;\|}{\sigma_{\min}\left(\;[\phi]^2 R[\psi]^{-2}R^t\;\right)} \leq \frac{\phi_\mu^2\;\|c\|}{\sigma_{\min}\left(\;[\phi]^2 R[\psi]^{-2}R^t\;\right)}. \tag{2.6}$$

In the example (2.3),

$$[\phi]^2 R[\psi]^{-2} = \begin{pmatrix} 0 & r_{12} & r_{13}\left(\frac{\psi_2}{\psi_3}\right)^2 & r_{14}\left(\frac{\psi_2}{\psi_4}\right)^2 & r_{15}\left(\frac{\psi_2}{\psi_5}\right)^2 & r_{16}\left(\frac{\psi_2}{\psi_6}\right)^2 \\ 0 & 0 & r_{23} & r_{24}\left(\frac{\psi_3}{\psi_4}\right)^2 & r_{25}\left(\frac{\psi_3}{\psi_5}\right)^2 & r_{26}\left(\frac{\psi_3}{\psi_6}\right)^2 \\ 0 & 0 & 0 & 0 & r_{35} & r_{26}\left(\frac{\psi_5}{\psi_6}\right)^2 \end{pmatrix}.$$

We now use the following lemma:

Lemma 2 *If $R \in \mathbb{R}^{n \times m}$ is an upper triangular matrix with rank n, then there exists $K_R > 0$ such that if $\psi \in \mathbb{R}^m$, $\psi > 0$, is sorted in increasing order and $\phi \in \mathbb{R}^n$ is defined by $\phi_i = \psi_{j(R,i)}$ then $\sigma_{\min}\left(\;[\phi]^2 R[\psi]^{-2}R^t\;\right) \geq K_R$.*

There are $m!$ possible permutations of m columns. Therefore, there are at most $m!$ matrices Q and R that can be generated from A and there exists $k_\Pi > 0$ such that $K_{R,\eta} \geq k_\Pi$ for all possible R's and Q's. From (2.5) and (2.6), we get

$$\|d\| \leq \frac{\phi_\mu^2}{k_R}\|c\| \leq \frac{\phi_\mu^2\,\|c\|}{k_\Pi} \leq \frac{\|c\|\,\|A\|_F^2}{k_\Pi\,\eta^2}\,c^t d \leq \frac{\|c\|\,\|A\|_F^2}{k_\Pi^3}\,c^t d.$$

To complete the proof of Lemma 1, take $K_\Pi = \|c\|\,\|A\|_F^2/k_\Pi^3$ •

Proof of Lemma 2. Let C be equal to $[\phi]^2 R[\psi]^{-2}$. Its entries are

$$c_{ij} = r_{ij}\left(\frac{\phi_i}{\psi_j}\right)^2. \tag{2.7}$$

For $j \geq j(R,1)$, let β_j be the biggest k such that $j(R,k) \leq j$. For $j < j(R,1)$, take $\beta_j = 0$. In the example (2.3), $\beta = (0,1,2,2,3,3)^t$. Making $\prod_{k=r}^{s} p_k = 1$ if $s < r$, and noticing that $r_{ij} = 0$ if $j < j(R,i)$, it follows from (2.7) that

$$c_{ij} = r_{ij}\left(\frac{\phi_{\beta_j}}{\psi_j}\right)^2 \prod_{k=i}^{\beta_j - 1}\left(\frac{\phi_k}{\phi_{(k+1)}}\right)^2.$$

For example, if R is given by (2.3), then $\beta_6 = 3$ and

$$c_{16} = r_{16}\left(\frac{\phi_1}{\psi_6}\right)^2 = r_{16}\left(\frac{\psi_2}{\psi_6}\right)^2 = r_{16}\left(\frac{\phi_3}{\psi_6}\right)^2\left(\frac{\phi_1}{\phi_2}\right)^2\left(\frac{\phi_2}{\phi_3}\right)^2 = r_{16}\left(\frac{\psi_5}{\psi_6}\right)^2\left(\frac{\psi_2}{\psi_3}\right)^2\left(\frac{\psi_3}{\psi_5}\right)^2.$$

Let us define $\Phi \in I\!R^n$ by $\Phi_n = 1$ and, for $k < n$,

$$\Phi_k = \left(\frac{\phi_k}{\phi_{(k+1)}}\right)^2. \tag{2.8}$$

Let $\Psi \in I\!R^m$ be given by $\Psi_j = 1$ for $j \leq j(R,1)$ and, for $j > j(R,1)$,

$$\Psi_j = \left(\frac{\phi_{\beta_j}}{\psi_j}\right)^2. \tag{2.9}$$

In the example (2.3),

$$\Phi = \left((\frac{\psi_2}{\psi_3})^2, (\frac{\psi_3}{\psi_5})^2, 1\right) \quad \text{and} \quad \Psi = \left(1,1,1,\left(\frac{\psi_3}{\psi_4}\right)^2, 1, \left(\frac{\psi_5}{\psi_6}\right)^2\right).$$

Notice that $\Psi_{j(R,i)} = 1$, because $\beta(j(R,i)) = i$. Therefore, (Φ,Ψ) belongs to the set

$$\mathcal{S}(R) = \{(u,v) \in I\!R^{n+m} \text{ s.t. } \forall i, v_{j(R,i)} = u_n = 1 \text{ and } 0 \leq u_i, v_i \leq 1\}. \tag{2.10}$$

Let us define $F(R,u,v) \in \mathbb{R}^{n\times n}$, for $(u,v) \in \mathcal{S}(R)$, by

$$f_{ij}(R,u,v) = r_{ij}\ v_j \prod_{k=i}^{\beta_j-1} u_k. \tag{2.11}$$

Thus, $C = F(R,\Psi,\Phi)$.

In the example (2.3) we have

$$F(R,u,v) = \begin{pmatrix} 0 & r_{12} & r_{13}u_1 & r_{14}u_1v_4 & r_{15}v_5u_1u_2 & r_{16}v_6u_1u_2 \\ 0 & 0 & r_{23} & r_{24}v_4 & r_{25}v_5u_2 & r_{26}v_6u_2 \\ 0 & 0 & 0 & 0 & r_{35} & r_{26}v_6 \end{pmatrix} \tag{2.12}$$

The set $\mathcal{S}(R)$ is compact and $\sigma_{\min}(F(R,u,v)R^t)$ depends continuously on (u,v). Therefore, $\sigma_{\min}(F(R,u,v)R^t)$ has a minimum $(\bar{u},\bar{v})$ in $\mathcal{S}(R)$. If we take $K_R = \sigma_{\min}(F(\bar{u},\bar{v})R^t)$, then $\sigma_{\min}(CR^t) \geq K_R$. We have then the following lemma

Lemma 3. *If $R \in \mathbb{R}^{n\times m}$ is an upper triangular matrix with rank n and $(u,v) \in \mathcal{S}(R)$, then $F(R,u,v)R^t$ is nonsingular.*

It follows from Lemma 3 that $K_R > 0$ and the proof of Lemma 2 is complete •

Proof of Lemma 3. We use induction on $nz(u)$, the number of zero entries of u. Let us start with the case $nz(u) = 0$. We can assume that $v > 0$. Otherwise, if $v_j = 0$ for some j's then $j \neq j(R,i)$ for all i, since $v_{j(R,i)} = 1$. We drop the corresponding columns of R and entries of v and get R', with rank n, and $v' > 0$ such that $F(R,u,v)R^t = F(R',u,v')(R')^t$.

The idea now is to bring back the ϕ's and ψ's. Take $\phi_n = 1$ and use use (2.8) to find ϕ_k for $k < n$: $\phi_k = \sqrt{u_k}\phi_{k+1}$. Then use (2.9) to define ψ: $\psi_j = 1$ for $j < j(R,1)$ and $\psi_j = \sqrt{v_j}/\phi_{\beta_j}$ otherwise. We have that $F(R,u,v)R^t = [\phi]^2 R[\psi]^{-2}$ and $F(R,u,v)R^t$ is nonsingular, because R has rank n and $\phi,\psi > 0$. Thus, we are done with $nz(u) = 0$.

Suppose now that Lemma 3 is true if $nz(u) < k$ and assume that $nz(u) = k$. Let u_s be one of the zero components of u. We must have $s < n$, because $u_n = 1$. Let us look at example (2.12) with $s = 2$. Notice that u_2 appears in F_{15}, F_{16}, F_{25} and F_{26}. Therefore,

$$F(R,u,v) = \begin{pmatrix} 0 & r_{12} & r_{13}u_1 & r_{14}u_1v_4 & 0 & 0 \\ 0 & 0 & r_{23} & r_{24}v_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & r_{35} & r_{26}v_6 \end{pmatrix}$$

and $F(R,u,v)R^t$ has the form

$$F(R,u,v)R^t = \begin{pmatrix} F(U,u_-,v_-)\,U^t & 0 \\ F(W,u_+,v_+)\,V^t & F(W,u_+,v_+)\,W^t. \end{pmatrix},$$

where $u_- = (u_1, u_2)^t$, $u_+ = (u_3) = (1)$, $v_- = (v_1, v_2, v_3, v_4)^t = (1, 1, 1, v_4)^t$, $v_+ = (v_5, v_6)^t = (1, v_6)^t$, and

$$U = \begin{pmatrix} 0 & r_{12} & r_{13} & r_{14} \\ 0 & 0 & r_{23} & r_{24} \end{pmatrix}, \qquad V = \begin{pmatrix} r_{15} & r_{16} \\ r_{25} & r_{26} \end{pmatrix}, \qquad W = \begin{pmatrix} r_{35} & r_{36} \end{pmatrix}.$$

By induction, $F(U, u_-, v_-)U^t$ and $F(W, u_+, v_+)W^t$ are nonsingular. Therefore, $F(R, u, v)R^t$ is nonsingular and we are done with the proof in this example.

In the general case, partition R as

$$R = \begin{pmatrix} U & V \\ 0 & W \end{pmatrix},$$

where U has s rows and $j(R, s+1) - 1$ columns. Partition F accordingly:

$$F(R, u, v) = \begin{pmatrix} X(R, u, v) & Y(R, u, v) \\ 0 & Z(R, u, v). \end{pmatrix}.$$

If $j \geq j(R, s+1)$ then $\beta_j \geq s+1$ and $\beta_j - 1 \geq s$. Therefore, if $i \leq s$ and $j \geq j(R, s+1)$, then the term u_s is present in the product in (2.11) and $f_{ij}(R, u, v) = 0$. This shows that $Y(R, u, v) = 0$ and

$$F(R, u, v)R^t = \begin{pmatrix} X(R, u, v)\ U^t & 0 \\ Z(R, u, v)\ V^t & Z(R, u, v)\ W^t. \end{pmatrix}.$$

We leave to the reader the proof that $j(U, i) = j(R, i)$ and $X(R, u, v) = F(W, u_-, v_-)$, for $v_- = (v_1, \ldots, v_{(s-1)}, 1)$ and $u_- = (u_1, \ldots, u_{(s-1)}, 1)^t$. Moreover, $(u_-, v_-) \in \mathcal{S}(U)$. Similarly, $Z(R, u, v) = F(U, v_+, u_+)$, for $v_+ = (v_{j(R,s+1)}, \ldots, v_n)^t$ and $u_+ = (u_{(s+1)}, \ldots, v_n)^t$, and $(u_+, v_+) \in \mathcal{S}(W)$. Since $nz(u_-)$ and $nz(u_+)$ are smaller than $nz(u)$, it follows from the induction hypothesis that $X(R, u, v)U^t$ and $Z(R, u, v)W^t$ are nonsingular. Therefore, $F(R, u, v)R^t$ is nonsingular and the proof of Lemma 3 is complete •

3. Convergence to the optimum

Proof of Theorem 2. Let $\bar{x} = \bar{x}(x^0, \lambda, \Pi) \in \partial\mathcal{F}_\Pi$ be the limit of the sequence x^k, let α be the number of zero components of $\bar{\xi} = A^t\bar{x} - b$, and let P be a permutation such that the first α components of $\bar{\psi} = P\bar{\xi}$ are equal to 0. Notice that

$$\bar{\psi} = P\bar{\xi} = P(A^t\bar{x} - b) = \left(AP^t\right)^t \bar{x} - Pb$$

and the first α columns of AP^t correspond to the binding restrictions at $\bar{x}$. Therefore, the first α columns of AP^t are linearly independent and there exists a matrix B such that

$$BAP^t = D = \begin{pmatrix} I_s & R \\ 0 & S \end{pmatrix}. \tag{3.13}$$

It follows that $A = B^{-1}DP$ and $d(\Pi, x) = B^t(D[\psi]^{-2}D^t)^{-1}Bc$, for $\psi = P\xi$. Since $\xi^{k+1} = \xi^k - \lambda/\chi^k A^t d^k$ and $P^t A = D^t(B^t)^{-1}$, we have that

$$\psi^{k+1} = \psi^k - \lambda\frac{1}{\chi^k}PA^t d^k = \psi^k - \lambda\frac{1}{\chi^k}D^t(D[\psi^k]^{-2}D^t)^{-1}Bc.$$

Let us partition ψ^k as $(\phi^k, \eta^k) \in I\!R^\alpha \times I\!R^{m-\alpha}$, and partition $(D[\psi^k]^{-2}D^t)^{-1}Bc$ as $(u^k, v^k) \in I\!R^\alpha \times I\!R^{m-\alpha}$. It follows that

$$\phi^{k+1} = \phi^k - \lambda\frac{1}{\chi^k}u^k, \tag{3.14}$$

$$\eta^{k+1} = \eta^k - \lambda\frac{1}{\chi^k}v^k. \tag{3.15}$$

Partitioning $Bc = (w^1, w^2)$ accordingly, we obtain

$$\left([\phi^k]^{-2} + R[\eta^k]^{-2}R^t\right)u^k + R[\eta^k]^{-2}S^t v_k = w^1, \tag{3.16}$$

$$S[\eta^k]^{-2}R^t u^k + S[\eta^k]^{-2}S^t v^k = w^2. \tag{3.17}$$

Let us make the change of variables $r^k = [\phi^k]^{-2}u^k$. The last equations become

$$\begin{pmatrix} r^k \\ v^k \end{pmatrix} = \begin{pmatrix} I + R[\eta^k]^{-2}R^t[\phi^k]^2 & R[\eta^k]^{-2}S^t \\ S[\eta^k]^{-2}R^t[\phi^k]^2 & S[\eta^k]^{-2}S^t \end{pmatrix}^{-1} \begin{pmatrix} w^1 \\ w^2 \end{pmatrix}.$$

Since $\phi^k \to 0$ and $\eta^k \to \bar{\eta} > 0$, it follows that (r^k, v^k) converges to

$$\begin{pmatrix} \bar{r} \\ \bar{v} \end{pmatrix} = \begin{pmatrix} I & R[\bar{\eta}]^{-2}S^t \\ 0 & S[\bar{\eta}]^{-2}S^t \end{pmatrix}^{-1} \begin{pmatrix} w^1 \\ w^2 \end{pmatrix} = \begin{pmatrix} w^1 - R[\bar{\eta}]^{-2}S^t\bar{v} \\ (S[\bar{\eta}]^{-2}S^t)^{-1}w^2 \end{pmatrix}. \tag{3.18}$$

This shows that $u^k = [\phi^k]^2 r^k$ converges to 0. It also shows that $[\phi^k]^{-1}u^k \to 0$ and $[\eta^k]^{-1}v^k \to [\bar{\eta}]^{-1}\bar{v}$. Therefore the sequence $\chi^k = \max\{\max([\phi^k]^{-1}u^k), \max([\eta^k]^{-1}v^k)\}$, converges to $\max\{0, \max([\bar{\eta}]^{-1}\bar{v})\}$. Since η^k converges, it follows from (3.15) that v^k/χ^k converges to zero. Since χ^k converges, this implies that $v^k \to 0$. Using (3.18) we conclude that $w^2 = 0$ and $\bar{r} = w^1$.

We claim that all the components of w^1 are bigger than or equal to 0. In fact, suppose that $w^1_i < 0$ for some $1 \leq i \leq \alpha$. It follows that $\bar{r}_i < 0$ and there exists k_0

such that $k > k_0 \Rightarrow r_i^k < 0$. Therefore, if $k > k_0$ then $u_i^k = [\phi]^2 r_i^k < 0$ and (3.14) shows that $\phi^{k+1} > \phi^k$. However, $\phi^k > 0$ and this implies that ϕ does not converge to 0. This is an absurd and all the components of w^1 must be bigger than or equal to 0.

Notice that

$$c = B^{-1}\begin{pmatrix} w^1 \\ w^2 \end{pmatrix} = B^{-1}\begin{pmatrix} I_s \\ 0 \end{pmatrix} w^1.$$

Equation (3.13) shows that $B^{-1}(I_s, 0)^t$ is formed by the columns of A corresponding to the binding restrictions at $\bar{x}$. Therefore, c is a positive linear combination of these columns and $\bar{x}$ is optimal. This concludes the proof of Theorem 2 •

4. Convergence for two variables

Proof of Theorem 3. Let $\bar{x} \in \partial\mathcal{F}_\Pi$ be the limit of x^k and let α be the number of binding restrictions at $\bar{x}$. Take a permutation P such that the first α columns of AP^t correspond to these α restrictions. Let B be the matrix formed by the first α columns of AP^t. The restrictions of Π, $A^t x \geq b$, are equivalent to $PA^t x \geq Pb$. The first α inequalities in this last expression give $B^t x \geq f$, where $f \in I\!R^\alpha$ is the vector with the first α components of Pb.

The proof proceeds by contradiction: suppose that $\bar{x}$ is not optimal. This implies that the linear program

$$\Pi_2: \quad \text{minimize} \quad v^t y \quad \text{subject to} \quad By = c \quad \text{and} \quad y \geq 0,$$

where $v = B^t(\bar{x} - x^0)$, has no feasible solution. The dual of Π_2 is

$$\Pi_2': \quad \text{maximize} \quad c^t x \quad \text{subject to} \quad B^t x \leq v.$$

The restrictions of Π_2' are satisfied by $x = \bar{x} - x^0$. Since Π_2 is infeasible, Π_2' is unbounded and there exists $h \in I\!R^2$ such that $c^t h > 0$ and $B^t h \leq v$. Let $p \neq 0$ be a vector orthogonal to c such that $(B^t p)_1 \geq 0$. Since $c^t h > 0$, h and p are linearly independent. The matrix

$$D = \begin{pmatrix} \frac{1}{c^t h} h^t \\ p^t \end{pmatrix}$$

is nonsingular and

$$DAP^t = E = \begin{pmatrix} r^t & u^t \\ s^t & v^t \end{pmatrix}, \quad \text{and} \quad Dc = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{4.19}$$

where $r = B^t h / c^t h = v / c^t h$, $s = B^t p$, and the vectors u and v are in $I\!R^{m-\alpha}$. Notice that $v = f - B^t x^0 < 0$, because $x^0 \in \mathcal{F}_\Pi^+$. Therefore, $r < 0$.

Take $y^k = (D^t)^{-1}x^k$ and $\bar{y} = (D^t)^{-1}\bar{x}$. Since $B^t x^k > f$,

$$y_1^k r + y_2^k s > f. \tag{4.20}$$

Analogously,

$$\bar{y}_1 r + \bar{y}_2 s = f. \tag{4.21}$$

Moreover, $z^k = c^t x^k = (Dc)^t (D^t)^{-1} x^k = y_1^k$. Since $z^k < z^{k+1}$, we have

$$y_1^k > y_1^{k+1} > \bar{y}_1. \tag{4.22}$$

It follows from (4.20) and (4.21) that $(y_1^k - \bar{y}_1)r + (y_2^k - \bar{y}_2)s > 0$ for all k. Therefore, $(y_2^k - \bar{y}_2)s > 0$, because $r < 0$ and $y_1^k - \bar{y}_1 > 0$. This implies that all the components of s have the same sign. By the way we have chosen p, $s_1 = (B^t p)_1 \geq 0$. Therefore, $s > 0$.

Let us analyze the evolution of $\psi^k = P\xi^k$ and y^k. Some algebra show that

$$\begin{aligned} y^{k+1} &= y^k - \frac{\lambda}{\chi^k}\Delta^k, && (4.23)\\ \psi^{k+1} &= \psi^k - \frac{\lambda}{\chi^k}E^t\Delta^k, && (4.24)\\ \chi^k &= \max\left([\psi^k]^{-1}E^t\Delta^k\right), && (4.25) \end{aligned}$$

where

$$\Delta^k = (D^t)^{-1}(D^{-1}EP[\xi^k]P^tE^t(D^t)^{-1})^{-1}c = (E[\psi^k]^{-2}E^t)^{-1}Dc. \tag{4.26}$$

Let us partition ψ^k as $(\phi^k, \eta^k)^t \in I\!R^\alpha \times I\!R^{m-\alpha}$. Multiplying (4.26) by $E[\psi^k]^{-2}E^t$ and using that $Dc = (1,0)^t$, we get

$$\begin{aligned} \left(r^t[\phi^k]^{-2}r + u^t[\eta^k]^{-2}u\right)\Delta_1^k + \left(r^t[\phi^k]^{-2}s + u^t[\eta^k]^{-2}v\right)\Delta_2^k &= 1, && (4.27)\\ \left(s^t[\phi^k]^{-2}r + v^t[\eta^k]^{-2}u\right)\Delta_1^k + \left(s^t[\phi^k]^{-2}s + v^t[\eta^k]^{-2}v\right)\Delta_2^k &= 0. && (4.28) \end{aligned}$$

Since $s, \phi^k, \eta^k > 0$, we can use (4.28) to conclude that

$$\Delta_2^k = -\frac{s^t[\phi^k]^{-2}r + v^t[\eta^k]^{-2}u}{s^t[\phi^k]^{-2}s + v^t[\eta^k]^{-2}v}\Delta_1^k \tag{4.29}$$

Equations (4.19) and (4.24) imply that

$$\phi_i^{k+1} = \phi_i^k - \frac{\lambda}{\chi^k}\left(\Delta_1^k r_i + \Delta_2^k s_i\right) = \phi_i^k - \frac{\lambda\Delta_1 s_i}{\chi^k}\left(\frac{r_i}{s_i} - \frac{s^t[\phi^k]^{-2}r + v^t[\eta^k]^{-2}u}{s^t[\phi^k]^{-2}s + v^t[\eta^k]^{-2}v}\right).$$

It follows from (4.22) and (4.23) that $\lambda\Delta_1/\chi^k > 0$. Since $s_i > 0$, $\phi_i^k - \phi_i^{k+1}$ has the same sign as

$$\frac{r_i}{s_i} - \frac{s^t[\phi^k]^{-2}r + v^t[\eta^k]^{-2}u}{s^t[\phi^k]^{-2}s + v^t[\eta^k]^{-2}v} = \frac{\frac{r_i}{s_i}\sum_{j=1}^{\alpha}\frac{s_j^2}{(\phi_j^k)^2} + \frac{r_i}{s_i}v^t[\eta^k]^{-2}v - \sum_{j=1}^{\alpha}\frac{r_j s_j}{(\phi_i^k)^2} + v^t[\eta^k]^{-2}u}{s^t[\phi^k]^{-2}s + v^t[\eta^k]^{-2}v},$$

which has the same sign as

$$\sum_{j=1}^{\alpha}\left(\frac{r_i}{s_i} - \frac{r_j}{s_j}\right)\left(\frac{s_j}{\phi_j^k}\right)^2 + \frac{r_i}{s_i}v^t[\eta^k]^{-2}v + v^t[\eta^k]^{-2}u.$$

Since $\phi_i^k > 0$ and $\lim_{k\to\infty}\phi_i^k = 0$, for every $k_0 \in \mathbb{N}$ there exists $k > k_0$ such that $\phi_i^k - \phi_i^{k+1} > 0$. Therefore, for every $k_0 \in \mathbb{N}$, there exists $k > k_0$ such that

$$\sum_{j=1}^{\alpha}\left(\frac{r_i}{s_i} - \frac{r_j}{s_j}\right)\left(\frac{s_j}{\phi_j^k}\right)^2 + \frac{r_i}{s_i}v^t[\eta^k]^{-2}v + v^t[\eta^k]^{-2}u > 0.$$

Take ν such that $r_\nu/s_\nu = \min\{r_i/s_i,\ i = 1,\ldots,\alpha\}$. For all $k_0 \in \mathbb{N}$, there exists $k > k_0$ such that

$$\sum_{j=1}^{\alpha}\left(\frac{r_\nu}{s_\nu} - \frac{r_j}{s_j}\right)\left(\frac{s_j\mu^k}{\phi_j^k}\right)^2 > -\left(\mu^k\right)^2\left(\frac{r_\nu}{s_\nu}v^t[\eta^k]^{-2}v + v^t[\eta^k]^{-2}u\right), \tag{4.30}$$

where $\mu^k = \max\left(\phi^k\right)$. Since $\mu^k \to 0$ and $\eta^k \to \bar{\eta} > 0$, the left hand side of (4.30) converges to 0 and for every $\epsilon > 0$, there exist k such that

$$0 \geq \sum_{j=1}^{\alpha}\left(\frac{r_\nu}{s_\nu} - \frac{r_j}{s_j}\right)\left(\frac{s_j\mu^k}{\phi_j^k}\right)^2 > -\epsilon.$$

Therefore, $r_j/s_j = r_\nu/s_\nu$ for all j, because $s_j\mu^k/\phi_j^k \geq s_j > 0$. Thus, $r = \beta s$, for $\beta = r_\nu/s_\nu$.

Some algebra and (4.29) show that

$$[\psi^k]^{-1}E^t\Delta^k = \Delta_1\begin{pmatrix} [\phi^k]^{-1}\frac{\beta v^t[\eta^k]^{-2}v - v^t[\eta^k]^{-2}u}{s^t[\phi^k]^{-2}s + v^t[\eta^k]^{-2}v}s \\ [\eta^k]^{-1}\left(u - \frac{\Delta_2^k}{\Delta_1^k}\right)\end{pmatrix}. \tag{4.31}$$

Since $s > 0$, (4.31) and (4.29) imply that $\lim_{k\to\infty}[\psi^k]^{-1}E^t\Delta^k = (0, [\bar{\eta}]^{-1}(u - \beta v))^t$ and (4.25) shows that χ^k converges to some (finite) $\bar{\chi}$.

Solving (4.27) and (4.28) and using that $r = \beta s$, we get

$$\Delta_1^k = \frac{1 + \frac{v^t[\eta^k]^{-2}v}{s^t[\phi^k]^{-2}s}}{(u-\beta v)^t\,[\eta^k]^{-2}\,(u-\beta v) + \frac{u^t[\eta^k]^{-2}u\; v^t[\eta^k]2v - \left(v^t[\eta^k]^{-2}u\right)^2}{s^t[\phi^k]^{-2}s}}. \tag{4.32}$$

The matrix E in (4.19) has the same rank as A: 2. Since $r = \beta s$, we must have $u \neq \beta v$. Taking the limit in (4.32) we obtain

$$\lim_{k\to\infty} \Delta_1^k = \frac{1}{(u-\beta v)^t\,[\bar\eta]^{-2}\,(u-\beta v)} > 0.$$

This shows that $z^{k+1} - z^k = y_1^{k+1} - y_1^k = \lambda\Delta_1^k/\chi^k$ does not converge to zero, which is absurd since $z^k = c^t x^k$ converges to $c^t\bar{x}$. Therefore $\bar{x}$ must be optimal and the proof of Theorem 3 is complete •

References

[M] W. Mascarenhas, *The affine scaling algorithm fails for* $\lambda = 0.999$. Relatório de Pesquisa RP 61/93, oct. 1993, IMECC, Universidade Estadual de Campinas, Campinas, 13081-022, Brazil. submitted to SIAM Journal of Optimization.

[MTW] R.D.C. Monteiro, T. Tsuchiya, Y. Wang, *A Simplified Global Convergence Proof of the Affine Scaling Algorithm.* SIE Working paper 92-20. SIE Department, University of Arizona, Tucson, AZ 85721, 1992 (accepted in Annals of Operations Research).

[T] T. Tsuchiya, *Global convergence of the affine scaling method for primal degenerate linear programming problems.* Math. of Operations Research, vol 17, 3, 1992, pg. 527.

RELATÓRIOS DE PESQUISA — 1993

01/93 On the Convergence Rate of Spectral Approximation for the Equations for Nonhomogeneous Asymmetric Fluids — *José Luiz Boldrini* and *Marko Rojas-Medar.*

02/93 On Fraisse's Proof of Compactness — *Xavier Caicedo* and *A. M. Sette.*

03/93 Non Finite Axiomatizability of Finitely Generated Quasivarieties of Graphs — *Xavier Caicedo.*

04/93 Holomorphic Germs on Tsirelson's Space — *Jorge Mujica* and *Manuel Valdivia.*

05/93 Zitterbewegung and the Electromagnetic Field of the Electron — *Jayme Vaz Jr.* and *Waldyr A. Rodrigues Jr.*

06/93 A Geometrical Interpretation of the Equivalence of Dirac and Maxwell Equations — *Jayme Vaz Jr.* and *Waldyr A. Rodrigues Jr.*

07/93 The Uniform Closure of Convex Semi-Lattices — *João B. Prolla.*

08/93 Embedding of Level Continuous Fuzzy Sets and Applications — *Marko Rojas-Medar, Rodney C. Bassanezi* and *Heriberto Román-Flores.*

09/93 Spectral Galerkin Approximations for the Navier-Stokes Equations: Uniform in Time Error Estimates — *Marko A. Rojas-Medar* and *José Luiz Boldrini.*

10/93 Semigroup Actions on Homogeneous Spaces — *Luiz A. B. San Martin* and *Pedro A. Tonelli.*

11/93 Clifford Algebra Approach to the Barut-Zanghi Model as a Hamiltonian System — *Jayme Vaz Jr.* and *Waldyr A. Rodrigues Jr.*

12/93 Propagation of Scalar Waves in Layered Media — *Lúcio Tunes dos Santos* and *Martin Tygel.*

13/93 On the Convergence of the NMO-Power Series for a Horizontally Stratified Medium — *Martin Tygel.*

14/93 Convergence Rates in the Sobolev H^s-Norm of Approximations by Discrete Convolutions — *Sônia M. Gomes.*

15/93 On the Choice of the Space Dimension in Ill-Posed Problems — *Cristina Cunha.*

16/93 Elliptic Equations in R^2 with Non-linearities in the Critical Range — *D. G. de Figueiredo, O. H. Miyagaki* and *B. Ruf.*

17/93 Drug Kinetics and Drug Resistance in Optimal Chemotherapy — *M. I. S. Costa, J. L. Boldrini* and *R. C. Bassanezi.*

18/93 Chemotherapeutic Treatments Involving Drug Resistance and Level of Normal Cells as a Criterion of Toxicity — *M. I. S. Costa, J. L. Boldrini* and *R. C. Bassanezi.*

19/93 Bifurcation of Singularities of Reversible Systems — *Marco Antonio Teixeira.*

20/93 Sistemas Não Lineares e Fractais — *Lúcio Tunes dos Santos.*

21/93 New Integral Representation of the Solution of Schrödinger Equation with Arbitrary Potential — *Rodolfo L. Monaco* and *Waldyr A. Rodrigues Jr.*

22/93 The Module of Derivations of a Stanley–Reisner Ring — *Paulo Brumatti* and *Aron Simis.*

23/93 On the Convergence Rate of Spectral Approximation for the Equations for Chemical Active Fluid — *Marko Rojas-Medar* and *Sebastián A. Lorca.*

24/93 Sufficient Conditions for Minima of some Translation Invariant Functionals — *Orlando Lopes.*

25/93 A Constrained Minimization Problem with Integrals on the Entire Space — *Orlando Lopes.*

26/93 O Pensamento Reducionista na Filosofia de Niels Bohr — *José Emílio Maiorino.*

27/93 On the first curve of the Fučik spectrum of an elliptic operator — *D.G. de Figueiredo* and *J.-P. Gossez.*

28/93 Generalização dos Testes de Shirley e de House — *Belmer Garcia Negrillo.*

29/93 Compacidad y Compactificación en Teoría de Modelos — *J. C. Cifuentes.*

30/93 Global Strong Solutions of the Equations for the Motion of Nonhogeneous Incompressible Fluids — *José Luiz Boldrini* and *Marko Rojas-Medar.*

31/93 A Equação de Laplace no Universo de de-Sitter-Castelnuovo — *D. Gomes* e *E. Capelas de Oliveira.*

32/93 Klein-Gordon Wave Equation in the de Sitter Universe — *E. Capelas de Oliveira* and *E.A. Notte Cuello.*

33/93 Mittag-Leffler Methods in Analysis — *Jorge Mujica.*

34/93 The Initial Value Problem for a Generalized Boussinesq Model — *Sebastián A. Lorca* and *José Luiz Boldrini.*

35/93 Problemas Minimax e Aplicações — *José Mario Martínez, Lúcio Tunes dos Santos* e *Sandra Augusta Santos.*

36/93 An Extension of the Theory of Secant Preconditioners — *José Mario Martínez.*

37/93 Convergence Estimates for the Wavelet-Galerkin Method: Superconvergence at the Node Points — *Sônia M. Gomes.*

38/93 An Error Estimate Uniform in Time for Spectral Semi-Galerkin Approximations of the Nonhomogeneous Navier-Stokes Equations — *J. L. Boldrini* and *M. Rojas-Medar.*

39/93 More About the Time Analysis of Tunnelling Processes — *Vladislav S. Olkhovsky* and *Erasmo Recami.*

40/93 Zero-Point Anomaly — *José Alexandre Nogueira* and *Adolfo Maia Jr.*

41/93 On a System of Evolution Equations of Magnetohydrodynamic Type — *José Luiz Boldrini* and *Marko Rojas-Medar.*

42/93 **Generalized Zeldovich's Regularization of the Vacuum Energy** — *José Alexandre Nogueira* and *Adolfo Maia Jr.*

43/93 **Global Strong Solution of the Equations for the Motion of a Chemical Active Fluid** — *M.A. Rojas-Medar* and *S.A. Lorca.*

44/93 **A Theorem of Finiteness for Fat Bundles** — *Lucas M. Chaves.*

45/93 **Partículas Elementares como Micro-Universos** — *E. Recami, V. T. Zanchin* and *M. T. Vasconselos.*

46/93 **Micro–Universes and "Strong Black–Holes": A Purely Geometric Approach to Elementary Particles** — *E. Recami, F. Raciti, W.A. Rodrigues Jr.* and *V. T. Zanchin.*

47/93 **The Tolman "Antitelephone" Paradox: Its Solution by Tachyon Mechanics** — *Erasmo Recami.*

48/93 **Radial Symmetry of Minimizers for Some Translation and Rotation Invariant Functionals** — *Orlando Lopes.*

49/93 **A Riemann Integral Approach to Feynman's Path Integral** — *Rodolfo L. Monaco, Roberto E. Lagos* and *Waldyr A. Rodrigues Jr.*

50/93 **The Relationship between Electromagnetism and Quantum Mechanics, and a Non-Linear Dirac Equation** — *C. Daviau, Waldyr A. Rodrigues Jr.* and *Jayme Vaz Jr.*

51/93 **A New Approach for the JWKB Theory** — *R.L. Monaco* and *E. Capelas de Oliveira.*

52/93 **An Error Estimate Uniform in Time for Spectral Galerkin Approximations for the Equations for the Motion of a Chemical Active Fluid** — *Marko A. Rojas-Medar* and *Sebastián A. Lorca.*

53/93 **On R-Automorphisms of $R[X]$** — *Miguel Ferrero* and *Antonio Paques.*

54/93 **Triangular Decomposition Methods for Solving Reducible Nonlinear Systems of Equations** — *J. E. Dennis Jr., José Mario Martínez* and *Xiaodong Zhang.*

55/93 **A Note on Discontinuous Vector Fields and Reversible Mappings** — *Marco Antonio Teixeira.*

56/93 **Shock Formation for a System of Conservation Laws in Two Space Dimensions** — *M.C. Lopes-Filho* and *H.J. Nussenzveig Lopes.*

57/93 **Multidimensional Hyperbolic Systems with Degenerate Characteristic Structure** — *M.C. Lopes-Filho* and *H.J. Nussenzveig Lopes.*

58/93 **On the Topology of Complete Riemannian Manifolds with Nonnegative Curvature Operator** — *Francesco Mercuri* and *Maria Helena Noronha.*

59/93 **The Initial Value Problem for the Equations of Magnetohydrodynamic Type in Non-Cylindrical Domains** — *Marko A. Rojas-Medar* and *R. Beltrán-Barrios.*

60/93 **A Minimax Method with Application to the Initial Vector Coding Problem** — *José Mario Martínez, Lúcio Tunes dos Santos* and *Sandra Augusta Santos.*

61/93 **The Affine Scaling Algorithm Fails for $\lambda = 0.999$** — *Walter F. Mascarenhas.*

62/93 **Maximal Abelian Normal Subgroups of Galois Pro-2-Groups** — *Antonio José Engler* and *João Bosco Nogueira.*

63/93 The Mathematical Structure of Newtonian Spacetime Classical Dynamics and Gravitation — *Waldyr A. Rodrigues, Jr., Quintino A. G. de Souza* and *Yuri Bozhkov.*

64/93 Dirac-Hestenes Spinor Fields, Their Covariant Derivatives, and Their Formulation on Riemann-Cartan Manifolds — *Waldyr A. Rodrigues, Jr., Quintino A. G. de Souza, J. Vaz, Jr.* and *P. Lounesto.*

65/93 Semi-Lattices and Lattices of Bounded Continuous Functions — *João B. Prolla.*

66/93 Resolution of linear complementarity problems using minimization with simple bounds — *Ana Friedlander, José Mario Martínez* and *Sandra Augusta Santos.*

67/93 Inexact Newton Methods for Solving Nonsmooth Equations — *José Mario Martínez* and *Liqun Qi.*

68/93 On the Choquet-Deny Theorem for the Strict Topology — *João B. Prolla* and *Maria Sueli M. Roversi.*