

TECHNICAL RESEARCH REPORT

An Interior Point Method for Linear Programming, with n Active Set Flavor

by Andre L. Tits

T.R. 99-47



ISR develops, applies and teaches advanced methodologies of design and analysis to solve complex, hierarchical, heterogeneous and dynamic problems of engineering technology and systems for industry and government.

ISR is a permanent institute of the University of Maryland, within the Glenn L. Martin Institute of Technology/A. James Clark School of Engineering. It is a National Science Foundation Engineering Research Center.

Web site <http://www.isr.umd.edu>

An Interior Point Method for Linear Programming, with an Active Set Flavor

André L. Tits

Department of Electrical and Computer Engineering
and Institute for Systems Research
University of Maryland, College Park
MD 20742, USA

August 25, 1999

Abstract

It is now well established that, especially on large linear programming problems, the simplex method typically takes up a number of iterations considerably larger than recent interior-points methods in order to reach a solution. On the other hand, at each iteration, the size of the linear system of equations solved by the former can be significantly less than that of the linear system solved by the latter.

The algorithm proposed in this paper can be thought of as a compromise between the two extremes: conceptually an interior-point method, it ignores, at each iteration, all constraints except those in a small “active set” (in the dual framework). For sake of simplicity, in this first attempt, an affine scaling algorithm is used and strong assumptions are made on the problem. Global and local quadratic convergence is proved.

1 Introduction and Algorithm Statement

Consider the problem the linear programming problem (in dual form)

$$(P) \quad \text{minimize } \langle c, x \rangle \text{ s.t. } Ax \leq b, \quad x \in \mathbb{R}^n,$$

with A an $m \times n$ real matrix and c a non-zero real vector of dimension n . Let S denote the feasible set

$$S = \{x : Ax \leq b\}, \tag{1}$$

a (possibly unbounded) polytope. In many problems of interest, m is much larger than n , i.e., S has a large number of vertices. A simplex method applied to such problem would likely take up a large number of iterations, hopping from vertex to vertex. At each iteration it would solve a linear system of size n . On the other hand, an interior-point method aiming at solving the KKT equations

$$\begin{aligned} c + A^T \lambda &= 0 \\ \lambda_i (a_i x - b_i) &= 0 \quad i = 1, \dots, m, \end{aligned} \tag{2}$$

where λ is the m -vector of dual variables, would likely take up considerably fewer iterations, but would involve, at each iteration, the solution of a (structured) linear system of much larger size $n + m$. (See, e.g., [1] for background on interior-point methods.)

The idea investigated in this paper is as follows. Try to guess, at each iteration, as subset of the constraints, termed the “active set”, such that good progress can be made at that iteration even when all other constraints are ignored. If all ignored constraints are inactive at the solution of (P), than in principle nothing will be lost. Identifying these however is obviously not possible in practice, at least far from the solution, so if the active set is small it likely misses many of the active constraints at the solution. If the active set is cleverly selected though, there may be hope that enough progress will still be made at every iteration to drive the iterates to the neighborhood of the solution, where fast local convergence can then take place.

In this paper, we show that the idea just put forth holds some promise. For sake of relative simplicity of the analysis, we consider an interior-point method of the affine scaling variety and we make strong nondegeneracy assumptions. We select as the active set at a given iteration the n constraints closest to be active at the corresponding iterate. We borrow primal and dual stepsize rules from [2] (itself inspired from [3]) and closely follow the convergence analysis carried out in that paper. We show that the resulting algorithm is globally and locally quadratically convergent.

To proceed, let $\mathbf{m} = \{1, \dots, m\}$, and, for $i \in \mathbf{m}$, let a_i be the i th row of A , let b_i be the i th entry of b , and let

$$g_i(x) = a_i x - b_i.$$

Further, let

$$G(x) = \text{diag}(g_1(x), \dots, g_m(x)).$$

and, given any $\lambda \in \mathbb{R}^m$, let

$$\Lambda = \text{diag}(\lambda_1, \dots, \lambda_m).$$

Next, given $\Omega \subseteq \mathbf{m}$, say $\Omega = \{i_1, \dots, i_\ell\}$, where the i_j s are listed in increasing order, define

$$A_\Omega = [a_{i_1}^T, \dots, a_{i_\ell}^T]^T,$$

$$G_\Omega(x) = \text{diag}(g_{i_1}(x), \dots, g_{i_\ell}(x)),$$

and

$$\Lambda_\Omega = \text{diag}(\lambda_{i_1}, \dots, \lambda_{i_\ell}).$$

The feasible set S is given by (1), the strictly feasible S^0 set by

$$S^0 = \{x \in \mathbb{R}^n : g_i(x) < 0 \quad \forall i \in \mathbf{m}\},$$

and the solution set S^* by

$$S^* = \{x^* \in S : \langle c, x^* \rangle \leq \langle c, x \rangle \quad \forall x \in S\}.$$

A point $x^* \in S$ is said to be *stationary* for (P) if there exists $\lambda^* \in \mathbb{R}^m$ such that

$$\begin{aligned} c + A^T \lambda^* &= 0 \\ \lambda_i^* g_i(x^*) &= 0 \quad \forall i \in \mathbf{m} \end{aligned} \tag{3}$$

(In particular, all vertices of S are stationary.) If furthermore $\lambda^* \geq 0$, then x^* is a *KKT point* for (P) , i.e., since (P) is convex, $x^* \in S^*$.

The following two assumptions are made throughout.

Assumption A1. $S^0 \neq \emptyset$.

Assumption A2. S^* is nonempty and bounded.

Assumption A2 implies that A has full column rank.

Assumption A2'. Let $\Omega \subseteq \{1, \dots, m\}$ satisfy $|\Omega| \geq n$. Then A_Ω has full column rank.

Given $x \in S$, we denote by $I(x)$ the index set of active constraints at x , i.e.

$$I(x) = \{i \in \mathbf{m} : g_i(x) = 0\}.$$

Assumption A3. For all $x \in S$, $\{a_i : i \in I(x)\}$ is a linearly independent set.

The following iteration is an “active set” version of one investigated in [2].

Iteration IPAS.

Parameters. $\beta \in (0, 1)$, $\lambda_{\max} > 0$.

Data. $x \in S^0$, $\lambda_i > 0 \forall i \in \mathbf{m}$, $\Omega \subseteq \mathbf{m}$, with $|\Omega| \geq n$.

Step 1. Solve

$$\begin{bmatrix} 0 & A_\Omega^T \\ \Lambda_\Omega A_\Omega & G_\Omega(x) \end{bmatrix} \begin{bmatrix} \Delta x \\ \tilde{\lambda}_\Omega \end{bmatrix} = \begin{bmatrix} -c \\ 0 \end{bmatrix} \quad (\mathcal{L}_\Omega(x, \lambda)).$$

For $i = 1, \dots, m$, set

$$\tilde{\lambda}_i = \begin{cases} (\tilde{\lambda}_\Omega)_j & \text{if } i \text{ is the } j\text{th smallest index in } \Omega, \\ 0 & \text{if } i \notin \Omega. \end{cases}$$

Step 2. Updates.

(i) Compute the largest feasible stepsize

$$\bar{t} = \begin{cases} \infty & \text{if } a_i \Delta x \leq 0 \quad \forall i \in \mathbf{m}, \\ \min\{(-g_i(x)/a_i \Delta x) : a_i \Delta x > 0, i \in \mathbf{m}\} & \text{otherwise.} \end{cases} \quad (4)$$

Set

$$\hat{t} = \min\{1, \max\{\beta \bar{t}, \bar{t} - \|\Delta x\|\}\}. \quad (5)$$

Set $x^+ = x + \hat{t} \Delta x$.

(ii) If $\tilde{\lambda}_i \leq g_i(x)$ for some $i \in \mathbf{m}$, set

$$\lambda_i^+ = \lambda_i^0, \quad \forall i \in \mathbf{m}. \quad (6)$$

Otherwise, set

$$\lambda_i^+ = \min\{\lambda_{\max}, \max\{\tilde{\lambda}_i, \|\Delta x\|^2\}\}, \quad \forall i \in \mathbf{m}. \quad (7)$$

(iii) Pick $\Omega^+ \supseteq \tilde{\Omega}$ with

$$\tilde{\Omega} = \text{indexes of } n \text{ least negative } g_i(x^+)$$

□

The following result, proved in the appendix, implies that Step 1 in Iteration IPAS is well defined.

Lemma 1.1 *Let $x \in S$, let $\lambda \in \mathbb{R}^m$ with $\lambda_i > 0$ for all $i \in \mathbf{m}$, and let $\Omega \subseteq \mathbf{m}$ satisfy $|\Omega| \geq n$. Then*

$$M_{\Omega}(x, \lambda) := \begin{bmatrix} 0 & A_{\Omega}^T \\ \Lambda A_{\Omega} & G_{\Omega}(x) \end{bmatrix}$$

is nonsingular.

The next result shows that Iteration IPAS can be repeated indefinitely.

Lemma 1.2 *Let $x \in S^0$, let $\lambda \in \mathbb{R}^m$ with $\lambda_i > 0$ for all i , and let $\Omega \subseteq \mathbf{m}$ satisfy $|\Omega| \geq n$. Then Iteration IPAS generates quantities with the following properties: $\Delta x \neq 0$, $\hat{t} > 0$, $x^+ \in S^0$, $\lambda_i^+ > 0$ for all $i \in \mathbf{m}$ and $|\Omega^+| \geq n$.*

Proof. Since $c \neq 0$, if $\Delta x = 0$ then the first block of equations in the linear system implies that $\tilde{\lambda}_{\Omega}$ is nonzero, while, since $x \in S^0$, the second block implies that it is zero, a contradiction. The other claims are immediate. □

We now consider sequences $\{x^k\}$, $\{\lambda^k\}$ and $\{\Omega^k\}$ obtained by repeated application of Iteration IPAS. Thus, for all k , $x^k \in S^0$, $\lambda_i^k > 0$ for all $i \in \mathbf{m}$, and $|\Omega^k| \geq n$. Auxiliary quantities computed at iteration k will bear superscript k as well.

For future use, we rewrite $\mathcal{L}_{\Omega^k}(x^k, \lambda^k)$ as

$$\sum_{i \in \Omega^k} \tilde{\lambda}_i^k a_i^T + c = 0, \tag{8}$$

$$\lambda_i^k a_i \Delta x^k + g_i(x^k) \tilde{\lambda}_i^k = 0 \quad \forall i \in \Omega^k. \tag{9}$$

We will prove convergence of the algorithm as stated. Note that Step 2(iii) leaves room for heuristic enlargement of the “active set” Ω .

2 Convergence Analysis

The analysis carried out in this section is strongly inspired from that in [2], itself inspired from [3]. It is provided *in extenso* for ease of reference.

2.1 Global Convergence

We show that, under Assumptions A1-A3, the sequence $\{x^k\}$ converges to S^* , the set of solution points.

First, in view of Lemma 1.2, $\hat{t}^k > 0$ for all k and Δx^k never vanishes, and thus the sequence $\{x^k\}$ generated by the algorithm never becomes constant. The next result implies that the values of the objective function and of all constraint functions with negative multiplier estimates $\tilde{\lambda}_i$ must decrease at every iteration.

Proposition 2.1 *Let $x \in S^0$, let $\lambda \in \mathbb{R}^m$ with $\lambda_i > 0$ for all $i \in \mathbf{m}$, let $\Omega \subseteq \mathbf{m}$ with $|\Omega| \geq n$ and let Δx and $\tilde{\lambda}$ solve $\mathcal{L}_\Omega(x, \lambda)$. Then*

$$\langle c, \Delta x \rangle = -\langle \tilde{\lambda}_\Omega, A_\Omega \Delta x \rangle < 0 \quad (10)$$

and

$$a_i \Delta x < 0 \quad \forall i \text{ s.t. } \tilde{\lambda}_i < 0. \quad (11)$$

Proof. See the appendix. □

Corollary. The sequence $\{x^k\}$ is bounded.

Proof. Assumption A2 implies that, given any $x^0 \in S$, the level set $\{x \in S : \langle c, x \rangle \leq \langle c, x^0 \rangle\}$ is bounded. The claim then follows from the monotone decrease of $\langle c, x^k \rangle$. □

We first show that $\{x^k\}$ converges to stationary points of (P) . The proofs of Lemmas 3.2 and 3.3 are given in the appendix.

Lemma 2.2 *Let $x^* \in \mathbb{R}^n$ and suppose that K , an infinite index set, is such that $\{x^k\}$ converges to x^* on K . If $\{\Delta x^k\}$ converges to zero on K , then x^* is stationary and $\{\lambda^k\}$ converges to λ^* on K , where λ^* is the unique multiplier vector associated with x^* .*

Lemma 2.3 *Let $x^* \notin S^*$ and suppose that K , an infinite index set, is such that $\{x^k\}$ converges to x^* on K . Then $\{\Delta x^k\}$ goes to zero on K .*

Proposition 2.4 $\{x^k\}$ converges to the set of stationary points of (P) .

Proof. By contradiction. Suppose not. Then, since $\{x^k\}$ is bounded, there exists some infinite index set K and some x^* not stationary such that $x^k \rightarrow x^*$ as $k \rightarrow \infty$, $k \in K$. In view of Lemma 2.2, $\{\Delta x^k\}$ does not converge to zero on K . Thus there exists an infinite index set $K' \subset K$ such that $\inf_{k \in K'} \|\Delta x^k\| > 0$. Since $x^k \rightarrow x^*$ as $k \rightarrow \infty$, $k \in K'$, this contradicts Lemma 2.3. Thus the claim holds. \square

Now note that, since $\langle c, x^k \rangle$ decreases at each iteration, if one limit point of $\{x^k\}$ is in S^* , then all of them are. Proceeding by contradiction towards proving that $\{x^k\}$ converges to S^* , we will assume that $\{x^k\}$ is bounded away from S^* .

Lemma 2.5 *If $\{x^k\}$ is bounded away from S^* , then $\{\Delta x^k\} \rightarrow 0$.*

Proof. By contradiction. Suppose there exists an infinite index set K such that $\inf_K \|\Delta x^k\| > 0$. Let $K' \subset K$, $x^* \in S$ be such that $x^k \rightarrow x^*$ as $k \rightarrow \infty$, $k \in K'$, with K' an infinite index set. Since $\{x^k\}$ is bounded away from S^* , it follows that $x^* \notin S^*$ which, in view of Lemma 2.3, leads to a contradiction. \square

The following key lemma is proved in the appendix.

Lemma 2.6 *Suppose $\{x^k\}$ is bounded away from S^* . Let x^* and x^{l*} be two limit points of $\{x^k\}$. In view of Proposition 2.4, they are stationary points. Let λ^* and λ^{l*} be the associated multiplier vectors. Then $\lambda^* = \lambda^{l*}$.*

Theorem 2.7 $\{x^k\}$ converges to S^* .

Proof. Proceeding again by contradiction, suppose that some limit point of $\{x^k\}$ is not in S^* and thus, since $\langle c, x \rangle$ takes on the same value at all limit points of $\{x^k\}$, that $\{x^k\}$ is bounded away from S^* . In view of Lemma 2.5, $\{\Delta x^k\} \rightarrow 0$. Let λ^* be the common multiplier vector associated with all limit points of $\{x^k\}$ (see Lemma 2.6). A simple contradiction argument shows that Lemma 2.2 then implies that $\{\tilde{\lambda}^k\} \rightarrow \lambda^*$. Since $\{x^k\}$ is bounded away from

S^* , $\lambda^* \not\geq 0$. Let i_0 be such that $\lambda_{i_0}^* < 0$. Then $\tilde{\lambda}_{i_0}^k < 0$ for all k large enough and thus $i_0 \in \Omega_k$ for all k large enough. Proposition 2.1 and the fact that $\hat{t}^k > 0$ for all k (Lemma 1.2) then imply that, for k large enough,

$$0 > g_{i_0}(x^k) > g_{i_0}(x^{k+1}) > \dots$$

contradicting the fact that $\{g_{i_0}(x^k)\} \rightarrow 0$. □

2.2 Local Rate of Convergence

Let x^* be a limit point of $\{x^k\}$ and let λ^* be the corresponding KKT multiplier vector. We now assume that the second order sufficiency conditions of optimality with strict complementarity holds at x^* , i.e.,

Assumption A4. $a_i v = 0 \forall i \in I(x^*)$ only if $v = 0$.

Assumption A5. $\lambda_i^* > 0$ for all $i \in I(x^*)$.

Assumptions A4 and A5 ensure that x^* is the unique solution of (P) and that exactly n constraints are active at x^* (i.e., x^* is a vertex). Since $\{x^k\}$ converges to S^* it follows that $\{x^k\} \rightarrow x^*$. The following result is a variation on a result of Fiacco and McCormick (proof of Theorem 14 in [4]) It is related to Lemma 2.1 (but Assumptions A4 and A5 are not in force in that lemma). For the sake of completeness, a proof is given in the appendix.

Proposition 2.8 *Let $\eta_i^* = \min\{\lambda_i^*, \lambda_{\max}\}$ and let $\Omega^* \supseteq I(x^*)$. Then the matrix $M_{\Omega^*}(x^*, \eta^*)$ is nonsingular.*

We show that, if $\lambda_i^* \leq \lambda_{\max}$ for all i , the pair $\{(x^k, \lambda^k)\}$ converges Q-quadratically to (x^*, λ^*) . First a preliminary result, also derived in [3] (again, it is proved in the appendix for ease of reference).

Lemma 2.9 *(i) $\{\Delta x^k\} \rightarrow 0$ and $\{\tilde{\lambda}^k\} \rightarrow \lambda^*$; (ii) for k large enough $\{i : \lambda_i^k > 0\} = I(x^*)$; (iii) if $\lambda_i^* \leq \lambda_{\max}$ for all $i \in \mathbf{m}$, then $\{\lambda^k\} \rightarrow \lambda^*$; (iv) for k large enough $\tilde{\Omega}^k = I(x^*)$.*

To prove Q-quadratic convergence of $\{(x^k, \lambda^k)\}$, the following property of Newton's method will be used. It is borrowed from [2]. Its proof is given in the appendix for ease of reference.

Proposition 2.10 *Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be twice continuously differentiable and let $z^* \in \mathbb{R}^n$ and $\rho > 0$ be such that $F(z^*) = 0$ and $\frac{\partial F}{\partial z}(z)$ is nonsingular whenever $z \in B(z^*, \rho) := \{z : \|z^* - z\| \leq \rho\}$. Let $d^N : B(z^*, \rho) \rightarrow \mathbb{R}^n$ be defined by $d^N(z) = -\left(\frac{\partial F}{\partial z}(z)\right)^{-1} F(z)$. Then given any $c_1 > 0$ there exists $c_2 > 0$ such that*

$$\|z^+ - z^*\| \leq c_2 \|z - z^*\|^2 \quad \forall z \in B(z^*, \rho) \quad (12)$$

for every $z \in B(z^*, \rho)$ and $z^+ \in \mathbb{R}^n$ for which, for each $i \in \{1, \dots, n\}$, either

$$(i) \quad |z_i^+ - z_i^*| \leq c_1 \|d^N(z)\|^2 \quad \forall z \in B(z^*, \rho)$$

or

$$(ii) \quad |z_i^+ - (z_i + d_i^N(z))| \leq c_1 \|d^N(z)\|^2 \quad \forall z \in B(z^*, \rho).$$

Theorem 2.11 *If $\lambda_i^* \leq \lambda_{\max} \quad \forall i \in \mathbf{m}$, then $\{(x^k, \lambda^k)\}$ converges to (x^*, λ^*) Q -quadratically.*

Proof. With reference to Proposition 2.10, let $\rho > 0$ be such that $M_{\Omega^k}(x, \lambda)$ is nonsingular for all $(x, \lambda) \in B((x^*, \lambda^*), \rho)$ and for all k large enough; in view of Proposition 2.8 and of the fact that Ω^k takes only finite many values, all of which contain $I(x^*)$, such ρ exists. Since $\{(x^k, \lambda^k)\} \rightarrow (x^*, \lambda^*)$ as $k \rightarrow \infty$, there exists k_0 such that $(x^k, \lambda^k) \in B((x^*, \lambda^*), \rho)$ for all $k \geq k_0$. Now let us first consider $\{\lambda^k\}$. For $i \in I(x^*)$, in view of the update rule for λ in Step 2(ii) of Iteration IPAS, $\lambda_i^{k+1} = \tilde{\lambda}_i^k$ for k large enough, so that condition (ii) in Proposition 2.10 holds for k large enough. For $i \notin I(x^*)$, for each k either again $\lambda_i^{k+1} = \tilde{\lambda}_i^k$ or $\lambda_i^{k+1} = \|\Delta x^k\|^2$. In the latter case, since $\lambda_i^* = 0$, condition (i) in Proposition 2.10 holds. Next, consider $\{x^k\}$. For $i \in \Omega^k \setminus I(x^*)$,

$$\frac{|\lambda_i^k|}{|\tilde{\lambda}_i^k|} = \frac{|g(x_i^k)|}{|a_i \Delta x^k|} \rightarrow \infty \quad \text{as } k \rightarrow \infty.$$

Thus,

$$\bar{t}^k = \min\{\lambda_i^k / \tilde{\lambda}_i^k : i \in I(x^*)\}$$

and

$$\hat{t}^k = \min\left\{1, \frac{\lambda_{i_k}^k}{\tilde{\lambda}_{i_k}^k} - \|\Delta x^k\|\right\} \quad (13)$$

for k large enough, for some $i_k \in I(x^*)$. In particular, t^k converges to 1. Thus, for k large enough and some $i_k \in I(x^*)$

$$\begin{aligned} \|x^{k+1} - (x^k + \Delta x^k)\| &= |\hat{t}^k - 1| \|\Delta x^k\| \\ &\leq \|\Delta x^k\| + \frac{|\tilde{\lambda}_i^k - \lambda^k|}{\tilde{\lambda}_i^k} \|\Delta x^k\|. \end{aligned}$$

Since $\lambda_i^* > 0$ for all $i \in I(x^*)$, it follows that for some $C > 1$ and all k large enough

$$\begin{aligned} \|x^{k+1} - (x^k + \Delta x^k)\| &\leq (\|\Delta x^k\| + C\|\tilde{\lambda}^k - \lambda^k\|)\|\Delta x^k\| \\ &\leq (1 + C)(\|\Delta x^k\| + \|\tilde{\lambda}^k - \lambda^k\|)^2 \end{aligned}$$

Thus condition (ii) of Proposition 2.10 holds. The claim then follows from Lemma 2.9 and Proposition 2.10. \square

3 Conclusion

An interior-point method has been proposed where only a small subset of the constraints (in the dual framework) is taken into account in each search direction computation. Global and local quadratic convergence has been proved. While an affine-scaling method has been used, and strong assumptions have been made, it is anticipated that similar ideas can be applied to more sophisticated interior-point methods and that the assumptions can be weakened, possibly subject to a more conservative selection of the active set. Finally, the update rule for the active set leaves room for heuristics (within the framework for which convergence has been proved) to speed up convergence in the early iterations.

4 Appendix: Some Proofs

Proof of Lemma 1.1. Let $(\Delta x, \tilde{\lambda}_\Omega)$ be such that $M_\Omega(x, \lambda) \begin{bmatrix} \Delta x \\ \tilde{\lambda}_\Omega \end{bmatrix} = 0$. Thus

$$A_\Omega^T \tilde{\lambda}_\Omega = 0 \tag{14}$$

$$\Lambda_\Omega A_\Omega \Delta x + G_\Omega(x) \tilde{\lambda}_\Omega = 0. \quad (15)$$

Taking the inner product of both sides of (14) by Δx yields

$$\langle \Delta x, A_\Omega^T \tilde{\lambda}_\Omega \rangle = 0. \quad (16)$$

Since $\lambda_i > 0$ for all $i \in \mathbf{m}$, left multiplying both sides of (15) by Λ_Ω^{-1} and taking the inner product with $\tilde{\lambda}_\Omega$ yields

$$\langle \tilde{\lambda}_\Omega, A_\Omega \Delta x \rangle + \langle \tilde{\lambda}_\Omega, \Lambda_\Omega^{-1} G_\Omega(x) \tilde{\lambda}_\Omega \rangle = 0. \quad (17)$$

From (16) and (17) we get

$$\langle \tilde{\lambda}_\Omega, \Lambda_\Omega^{-1} G_\Omega(x) \tilde{\lambda}_\Omega \rangle = 0.$$

Since the $\Lambda_\Omega^{-1} G_\Omega(x)$ is negative semidefinite, it follows that $G_\Omega(x) \tilde{\lambda}_\Omega = 0$. In view of (15) this implies that $A_\Omega \Delta x = 0$. It then follows from Assumption A2' that $\Delta x = 0$. Finally, equation (14) together with Assumption A3 and the fact that $G_\Omega(x) \tilde{\lambda}_\Omega = 0$ implies that $\tilde{\lambda}_\Omega = 0$.

Proof of Proposition 2.1. The proof will make use of the following lemma.

Lemma A.1. Let $x \in S$, $\lambda \in \mathbb{R}^m$ such that $\lambda_i > 0$ for all $i \in \mathbf{m}$, and let $(\Delta x, \tilde{\lambda}_\Omega)$ satisfy

$$\Lambda_\Omega A_\Omega \Delta x + G_\Omega(x) \tilde{\lambda}_\Omega = 0. \quad (18)$$

Then (i) $\langle \tilde{\lambda}_\Omega, A_\Omega \Delta x \rangle \geq 0$ and (ii) if $\Delta x \neq 0$, then $\langle \tilde{\lambda}_\Omega, A_\Omega \Delta x \rangle > 0$.

Proof. Left multiplying both sides of (18) by Λ_Ω^{-1} yields

$$A_\Omega \Delta x + \Lambda_\Omega^{-1} G_\Omega(x) \tilde{\lambda}_\Omega = 0 \quad (19)$$

and taking the inner product with $\tilde{\lambda}_\Omega$ yields

$$\langle \tilde{\lambda}_\Omega, A_\Omega \Delta x \rangle + \langle \tilde{\lambda}_\Omega, \Lambda_\Omega^{-1} G_\Omega(x) \tilde{\lambda}_\Omega \rangle = 0. \quad (20)$$

The first claim follows from negative semidefiniteness of $G_\Omega(x)$. Concerning the second claim, assume by contradiction that $\langle \tilde{\lambda}_\Omega, A_\Omega \Delta x \rangle = 0$. It follows from (20) that $G_\Omega(x) \tilde{\lambda}_\Omega = 0$. This, together with (19) and Assumption 2' implies that $\Delta x = 0$, proving the claim.

Proof of Proposition 2.1. Since, in view of Lemma 1.2, $\Delta x \neq 0$, the first claim directly follows from Lemma A.1. Since $x \in S^0$ and $\lambda_i > 0$ for all $i \in \Omega$, it follows from (18) that $a_i \Delta x < 0$ for all $i \in \Omega$ such that $\tilde{\lambda}_i < 0$, proving the second claim.

Proof of Lemma 2.2. Suppose $\{\Delta x^k\} \rightarrow 0$ as $k \rightarrow \infty$, $k \in K$. Without loss of generality (by going down to a further subsequence if necessary) assume that, for some $\hat{\Omega}$, $\Omega_k = \hat{\Omega}$ for all $k \in K$. Equation (8)-(9) then yield

$$\sum_{i \in \hat{\Omega}} \tilde{\lambda}_i^k a_i^T + c = 0, \quad (21)$$

and

$$\lambda_i^k a_i \Delta x^k + g_i(x^k) \tilde{\lambda}_i^k = 0 \quad \forall i \in \hat{\Omega}, \forall k. \quad (22)$$

Since $\{\lambda^k\}$ is bounded (by construction), it follows from (22) that for all $i \in \hat{\Omega}$ for which $g_i(x^*) < 0$, $\{\tilde{\lambda}_i^k\} \rightarrow 0$ as $k \rightarrow \infty$, $k \in K$. In view of (21) and of Assumption A3, it follows that, for all $i \in \hat{\Omega}$, $\{\tilde{\lambda}_i^k\}$ converges on K , say to λ_i^* . Letting $\lambda_i^* = 0$ for all $i \notin \hat{\Omega}$ and taking limits in (21)-(22) then yields

$$A^T \lambda^* + c = 0,$$

$$\lambda_i^* g_i(x^*) = 0, \quad i = 1, \dots, m,$$

implying that x^* is stationary, with multiplier vector λ^* .

Proof of Lemma 2.3. Let $J^k = \{i \in \mathbf{m} : \lambda_i^k \leq g_i(x^k)\}$. The proof will make use of the following result.

Lemma A.2. Let K be an infinite index set such that, on K , λ^k is bounded away from zero, $\tilde{\lambda}^k$ is bounded, Δx^k is bounded and, for some x^* , $x^k \rightarrow x^*$ as $k \rightarrow \infty$, $k \in K$. Then \bar{t}^k is bounded away from 0 on K .

Proof. According to Step 2(i) of Iteration IPAS, $\bar{t}^k = \min t_i^k$ where, for $i = 1, \dots, m$,

$$t_i^k = -g_i(x^k)/a_i \Delta x^k \text{ if } a_i \Delta x^k > 0 \quad (23)$$

and $t_i^k = \infty$ otherwise. Proceeding by contradiction suppose that, for some infinite index set $K' \subseteq K$ and some $i_0 \in \mathbf{m}$,

$$t_{i_0}^k \rightarrow 0 \text{ as } k \rightarrow \infty, \quad k \in K'. \quad (24)$$

Clearly,

$$t_{i_0}^k = -g_{i_0}(x^k)/a_{i_0} \Delta x^k, \quad \forall k \in K', k \text{ large enough.} \quad (25)$$

First suppose that i_0 belongs to Ω^k for infinitely many $k \in K'$. For all such k , in view of (9), $t_{i_0}^k = \lambda_{i_0}^k / \tilde{\lambda}_{i_0}^k$. Since $\tilde{\lambda}^k$ is bounded away from zero and λ^k is bounded, this contradicts (24). Thus $i_0 \notin \Omega^k$ for all $k \in K'$, k large enough. Since (25) is equivalently written as

$$g_{i_0}(x^k + t_{i_0}^k \Delta x^k) = 0, \quad \forall k \in K', k \text{ large enough,}$$

letting $k \rightarrow \infty$, $k \in K'$, we see that $g_{i_0}(x^*) = 0$. Now, in view of the update rule for Ω^k , whenever $i_0 \notin \Omega^k$, at least n other constraints $g_i(x^k)$ are closer to zero than $g_{i_0}(x^k)$. Since the number of choices of n constraints is finite, there must exist an infinite index set $K'' \subseteq K'$ and indexes $i_1, \dots, i_n \in \mathbf{m}$ such that, for all $k \in K''$, $g_{i_0}(x^k) < g_{i_\ell}(x^k)$, $\ell = 1, \dots, n$. Letting $k \rightarrow \infty$ on K'' , we conclude that $g_{i_\ell}(x^*) = 0$, $i = 0, 1, \dots, n$, in contradiction with linear independence Assumption A3. This contradiction completes the proof.

Lemma A.3. Let K be an infinite index set such that

$$\inf\{\|\Delta x^{k-1}\| : k \in K, \tilde{\lambda}_i^k > g_i(x^k) \forall i \in \Omega^k\} > 0.$$

Then $\{\Delta x^k\} \rightarrow 0$ as $k \rightarrow \infty$, $k \in K$.

Proof. In view of (6) and (7), for all $i \in \mathbf{m}$, λ_i^k is bounded away from zero on K . Proceeding by contradiction, assume that, for some infinite index set $K' \subset K$, $\inf_{k \in K'} \|\Delta x^k\| > 0$. Since $\{x^k\}$ and $\{\lambda^k\}$ are bounded, we may assume, without loss of generality, that for some x^* and λ^* , with $\lambda_i^* > 0$ for all i , and some Ω^* with $|\Omega^*| \geq n$,

$$\begin{aligned} \{x^k\} &\rightarrow x^* & \text{as } k \rightarrow \infty, & k \in K' \\ \{\lambda^k\} &\rightarrow \lambda^* & \text{as } k \rightarrow \infty, & k \in K' \\ \Omega^k &= \Omega^* & \forall k \in K'. \end{aligned}$$

Since in view of Lemma 1.1 and Assumption A3, $M(x^*, \lambda^*)$ is nonsingular, it follows that, for some v^* and $\tilde{\lambda}^*$, with $v^* \neq 0$ (since $\inf_{k \in K'} \|\Delta x^k\| > 0$),

$$\begin{aligned} \{\Delta x^k\} &\rightarrow v^* & \text{as } k \rightarrow \infty, & k \in K', \\ \{\tilde{\lambda}^k\} &\rightarrow \tilde{\lambda}^* & \text{as } k \rightarrow \infty, & k \in K'. \end{aligned}$$

In view of Lemma A.2, it follows that \bar{t}^k is bounded away from zero on K' , and so is \hat{t}^k (Step 2 (i) in Iteration IPAS), i.e., for some $\underline{t} > 0$, $\hat{t}^k \geq \underline{t}$ for all $k \in K'$. In view of Proposition 2.1 and Lemma A.1(i), it follows that

$$\langle c, x^{k+1} \rangle \leq \langle c, x^k \rangle - \underline{t} \langle \tilde{\lambda}^k, A \Delta x^k \rangle \quad \forall k \in K'.$$

Taking limits in (9) as $k \rightarrow \infty$, $k \in K'$, we get

$$\Lambda_{\Omega^*}^* A_{\Omega^*} v^* + G_{\Omega^*}(x^*) \tilde{\lambda}_{\Omega^*} = 0.$$

Since $v^* \neq 0$ and since $\tilde{\lambda}_i^* = 0$ for $i \notin \Omega^*$, it follows from Lemma A.1(ii) that $\langle \tilde{\lambda}^*, Av^* \rangle > 0$ and thus there exists $\delta > 0$ such that $\langle \tilde{\lambda}^k, A\Delta x^k \rangle > \delta$ for k large enough, $k \in K'$. Since, in view of Proposition 2.1, $\langle c, x^k \rangle$ is monotonic nonincreasing, it follows that $\langle c, x^k \rangle \rightarrow -\infty$ as $k \rightarrow \infty$, a contradiction since x^k is bounded.

Proof of Lemma 2.3. Let us again proceed by contradiction, i.e., suppose $\{\Delta x^k\}$ does not converge to zero as $k \rightarrow \infty$, $k \in K$. In view of Lemma A.3, there exists an infinite index set $K' \subset K$ such that

$$\tilde{\lambda}_i^{k-1} > g_i(x^{k-1}) \quad \forall i \in \Omega^{k-1}, \forall k \in K', \quad (26)$$

$$\Delta x^{k-1} \rightarrow 0 \text{ as } k \rightarrow \infty, k \in K'.$$

Also, without loss of generality, for some $\hat{\Omega}$, $\Omega^k = \hat{\Omega}$ for all $k \in K'$. Since $\{x^k\} \rightarrow x^*$ as $k \rightarrow \infty$, $k \in K$ and $\|x^k - x^{k-1}\| = \|\hat{t}^{k-1} \Delta x^{k-1}\| \leq \|\Delta x^{k-1}\|$, it follows that $\{x^{k-1}\} \rightarrow x^*$ as $k \rightarrow \infty$, $k \in K'$ which implies, in view of Lemma 2.2, that x^* is stationary and $\{\tilde{\lambda}^{k-1}\} \rightarrow \lambda^*$ as $k \rightarrow \infty$, $k \in K'$, where λ^* is the corresponding multiplier vector. From (26) it follows that $\lambda_i^* \geq 0$ for all $i \in \hat{\Omega}$ such that $g_i(x^*) = 0$. Since by construction of $\tilde{\lambda}$, $\lambda_i^* = 0$ for all $i \notin \hat{\Omega}$, it follows that all components of λ^* are nonnegative, thus that $x^* \in S^*$, a contradiction.

Proof of Lemma 2.6. Let L be the set of limit points of $\{x^k\}$ (in view of Proposition 2.4, all of these are stationary points of (P)). L is bounded (since $\{x^k\}$ is bounded) and, as a limit set, it is closed, thus compact. We first prove an auxiliary lemma.

Lemma A.4. If $\{x^k\}$ is bounded away from S^* , then L is connected.

Proof. Suppose not. Then there exists $E, F \subset \mathbb{R}^n$, both nonempty, such that $L = E \cup F$, $\overline{E} \cap F = \emptyset$, $E \cap \overline{F} = \emptyset$. Since L is compact E and F must be compact. Thus $\delta := \min_{x \in E, x' \in F} \|x - x'\| > 0$. A simple contradiction argument using the fact that $\{x^k\}$ is bounded shows that, for k large enough, $\min_{x \in L} \|x^k - x\| \leq \delta/3$, i.e., either $\min_{x \in E} \|x^k - x\| \leq \delta/3$ or $\min_{x \in F} \|x^k - x\| \leq \delta/3$. Moreover, since both E and F are nonempty (i.e., contain limit points of $\{x^k\}$), each of these situations occurs infinitely many times. Thus $K := \{k : \min_{x \in E} \|x^k - x\| \leq \delta/3, \min_{x \in F} \|x^{k+1} - x\| \leq \delta/3\}$ is an infinite

index set and $\|\Delta x^k\| \geq \delta/3 > 0$ for all $k \in K$. On the other hand since $\{x^k\}_{k \in K}$ is bounded and bounded away from S^* , it has some limit point $x^* \notin S^*$. In view of Lemma 2.3, this is a contradiction.

Proof of Lemma 2.6. Given any $x \in L$, let $\lambda(x)$ be the multiplier vector associated with x and let $J(x)$ be the index set of “binding” constraints at x , i.e.,

$$J(x) = \{i \in \mathbf{m} : \lambda_i(x) \neq 0\}.$$

First note that, in view of linear independence Assumption A3, if $x, x' \in L$ are such that $J(x) = J(x')$, then $\lambda(x) = \lambda(x')$. To conclude the proof, we show that, for any $x, x' \in L$, $J(x) = J(x')$. Let $\tilde{x} \in L$ be arbitrary and let $E := \{x \in L : J(x) = J(\tilde{x})\}$ and $F := \{x \in L : J(x) \neq J(\tilde{x})\}$. We show that both E and F are closed. Let $\{y^\ell\} \subseteq L$ be a convergent sequence, say to \hat{y} , such that $J(y^\ell) = J$ for all ℓ , for some J . It follows from the first part of this proof that $\lambda(y^\ell) = \lambda$ for all ℓ for some λ . Now, for all ℓ , $g_j(y^\ell) = 0$ for all j such that $\lambda_j \neq 0$, so that $g_j(\hat{y}) = 0$ for all j such that $\lambda_j \neq 0$. Thus $J \subseteq I(\hat{y})$ and from linear independence Assumption A3 it follows that $\lambda(\hat{y}) = \lambda$ and thus $J(\hat{y}) = J$. Also, since L is closed, $\hat{y} \in L$. Thus, if $\{y^\ell\} \subseteq E$ then $\hat{y} \in E$ and, if $\{y^\ell\} \subseteq F$ then $\hat{y} \in F$, proving that both E and F are closed. Since E is nonempty (it contains \tilde{x}), connectedness of L (Lemma A.4) implies that F is empty. Thus $J(x) = J(\tilde{x})$ for all $x \in L$, and the proof is complete.

Proof of Proposition 2.8. Let $(\Delta x, \tilde{\lambda}_{\Omega^*})$ be such that

$$M_{\Omega^*}(x^*, \eta^*) \begin{pmatrix} \Delta x \\ \tilde{\lambda}_{\Omega^*} \end{pmatrix} = 0.$$

Thus

$$A_{\Omega^*}^T \tilde{\lambda}_{\Omega^*} = 0 \tag{27}$$

$$\text{diag}(\eta_i^*) A_{\Omega^*} \Delta x + G_{\Omega^*}(x^*) \tilde{\lambda}_{\Omega^*} = 0. \tag{28}$$

In view of Assumption A5 and since $\Omega^* \supseteq I(x^*)$, (28) implies that

$$a_i \Delta x = 0 \quad \forall i \in I(x^*) \tag{29}$$

and, since $\eta_i^* = \lambda_i^* = 0$ when $i \notin I(x^*)$,

$$\tilde{\lambda}_i = 0 \quad \forall i \notin I(x^*). \tag{30}$$

In view of (29) and Assumption A4, it follows that $\Delta x = 0$. Finally, it follows from (27), (30) and Assumption A3 that $\tilde{\lambda}_{\Omega^*} = 0$.

Proof of Lemma 2.9. We first prove by contradiction that $\{\Delta x^k\} \rightarrow 0$. Thus suppose that there exists an infinite index set K such that $\inf_K \|\Delta x^k\| > 0$. In view of Lemma A.3, there exists an infinite index set $K' \subset K$ such that $\{\Delta x^{k-1}\} \rightarrow 0$ as $k \rightarrow \infty, k \in K'$ and, for all $k \in K'$, $\tilde{\lambda}_i^{k-1} > g_i(x^{k-1})$ for all $i \in \Omega^{k-1}$. Without loss of generality, assume that $\Omega^{k-1} = \Omega^*$ for all $k \in K'$, for some $\Omega^* \in \mathbf{m}$ with $|\Omega^*| \geq n$. It follows from Lemma 2.2 that $\{\tilde{\lambda}^{k-1}\} \rightarrow \lambda^*$ as $k \rightarrow \infty, k \in K'$ and from the update rule for λ^k in Step 2 (ii) of Iteration IPAS that, for all $i \in \mathbf{m}$, $\lambda_i^k \rightarrow \eta_i^* = \min\{\lambda_i^*, \lambda_{\max}\}$ as $k \rightarrow \infty, k \in K'$. Since for all k Ω^k contains the indexes of the n least negative $g_i(x^k)$ and since $x^k \rightarrow x^*$ and, from Assumption A3, $|I(x^*)| = n$, it follows that $\Omega^* \supseteq I(x^*)$. In view of Proposition 2.8 $M_{\Omega^*}(x^*, \eta^*)$ is nonsingular and thus, since $x^* \in S^*$, $\{\Delta x^k\} \rightarrow 0$ as $k \rightarrow \infty, k \in K'$, a contradiction. Thus $\{\Delta x^k\} \rightarrow 0$. It now follows from Lemma 2.2 that $\{\tilde{\lambda}^k\} \rightarrow \lambda^*$ and, in view of Assumption A5, that $\tilde{\lambda}_i^k > g_i(x^k)$ for all i for k large enough. The update rule for λ^k again implies that $\{\lambda_i^k\} \rightarrow \min\{\lambda_i^*, \lambda_{\max}\}$ for all $i \in \mathbf{m}$. Finally, since $|I(x^*)| = n$, it follows from Step 2(iii) in Iteration IPAS that $\tilde{\Omega}^k = I(x^*)$ for k large enough.

Proof of Proposition 2.10. First, let $i \in \{1, \dots, n\}$ be such that (i) holds. Since $\left(\frac{\partial F}{\partial z}(z)\right)^{-1}$ is bounded in $B(z^*, \rho)$ and F is Lipschitz continuous in the same ball, there exists $c_2 > 0$ such that, for all $z \in B(z^*, \rho)$

$$|z_i^+ - z_i^*| \leq c_1 \left\| \frac{\partial F}{\partial z}(z)^{-1} \right\|^2 \|F(z) - F(z^*)\|^2 \leq c_2 \|z - z^*\|^2.$$

Next, suppose (ii) holds. Then

$$\begin{aligned} |z_i^+ - z_i^*| &\leq |z_i^+ - (z + d^N(z))_i| + |z_i^* - (z + d^N(z))_i| \\ &\leq c_1 \|d^N(z)\|^2 + \|z^* - (z + d^N(z))\| \\ &\leq c_1 \|d^N(z)\|^2 + \left\| \frac{\partial F}{\partial z}(z)^{-1} \right\| \|F(z) + \frac{\partial F}{\partial z}(z)(z^* - z)\|. \end{aligned}$$

The first term in the right hand side is as in (i). Boundedness of $\frac{\partial F}{\partial z}(z)^{-1}$ in $B(z^*, \rho)$ and regularity of F thus again imply that the claim holds.

References

- [1] S.J. Wright. *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia, 1997.

- [2] A.L. Tits and J.L. Zhou. A simple, quadratically convergent algorithm for linear and convex quadratic programming. In W.W. Hager, D.W. Hearn, and P.M. Pardalos, editors, *Large Scale Optimization: State of the Art*, pages 411–427. Kluwer Academic Publishers, 1994.
- [3] E.R. Panier, A.L. Tits, and J.N. Herskovits. A QP-free, globally convergent, locally superlinearly convergent algorithm for inequality constrained optimization. *SIAM J. Contr. and Optim.*, 26(4):788–811, July 1988.
- [4] A.V. Fiacco and G.P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. Wiley, New-York, 1968.