Kybernetika

Christian Grossmann; Diethard Klatte; Bernd Kummer

Convergence of primal-dual solutions for the nonconvex log-barrier method without LICQ

Kybernetika, Vol. 40 (2004), No. 5, [571]--584

Persistent URL: `http://dml.cz/dmlcz/135617`

KYBERNETIKA — VOLUME 40 (2004), NUMBER 5, PAGES 571 – 584

CONVERGENCE OF PRIMAL–DUAL SOLUTIONS FOR THE NONCONVEX LOG–BARRIER METHOD WITHOUT LICQ

Christian Grossmann,* Diethard Klatte and Bernd Kummer

This paper is dedicated to our colleague, friend and teacher František Nožička on the occasion of his 85th birthday.

This paper characterizes completely the behavior of the logarithmic barrier method under a standard second order condition, strict (multivalued) complementarity and MFCQ at a local minimizer. We present direct proofs, based on certain key estimates and few well-known facts on linear and parametric programming, in order to verify existence and Lipschitzian convergence of local primal-dual solutions without applying additionally technical tools arising from Newton–techniques.

Keywords: log-barrier method, Mangasarian–Fromovitz constraint qualification, convergence of primal-dual solutions, locally linearized problems, Lipschitz estimates

AMS Subject Classification: 90C30, 65K10, 49K40, 49M37

1. INTRODUCTION

We consider the nonlinear programming problem

$$\begin{aligned} &f(x) \to \min ! \\ \text{subject to} \quad &x \in G = \{ x \in \mathbb{R}^n \,:\, g_i(x) \le 0,\ i = 1,\dots,m \}, \end{aligned} \tag{1}$$

where the functions f, $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1,\dots,m$ are twice continuously differentiable. In the log-barrier method, the original problem (1) is embedded into a family of auxiliary problems with positive embedding parameter s,

$$\begin{aligned} F(x,s) = f(x) - s \sum_{i=1}^{m} \ln(-g_i(x)) &\to \min ! \\ \text{subject to} \quad x \in G^0 = \{ x \in \mathbb{R}^n \,:\, g_i(x) < 0,\ i = 1,\dots,m \}&. \end{aligned} \tag{2}$$

*This author was supported by DFG grant GR 1777/2–2.

For any local solution $x(s)$ of problem (2), the necessary optimality condition yields

$$\nabla f(x(s)) - \sum_{i=1}^{m} \frac{s}{g_i(x(s))} \nabla g_i(x(s)) = 0. \tag{3}$$

Hence, with $L(x,y) = f(x) + \sum_{i=1}^{m} y_i g_i(x)$, the Lagrange condition

$$D_x L(x(s), y(s)) = 0 \tag{4}$$

holds with the so-called *log-barrier multipliers* defined by

$$y_i(s) = -\frac{s}{g_i(x(s))}, \quad i = 1, \ldots, m. \tag{5}$$

Given a stationary solution x^* of (1), suitable second order conditions yield that x^* is an isolated local minimizer. If, in addition, the linear independence constraint qualification (LICQ) and the strict complementarity condition are fulfilled, then, by a classical result (cf. Fiacco and McCormick [4, Thm. 14]), a differentiable local primal-dual path $(x(s), y(s))$ leading to (x^*, y^*) is generated by the log-barrier method as well as by several further barrier-penalty methods (cf. [4]). Under the same assumptions, this result can be extended to rather wide classes of barrier-penalty methods (cf., e.g., [6, 7, 13]).

However, if LICQ does not hold, the associated dual solution is not uniquely defined and is hence often called *degenerate*. In the literature of the last decade, there is a growing interest in the convergence behavior of numerical methods in the case of degenerate solutions. As examples let us refer to recent papers [5, 16, 18, 19, 20] on interior methods and modified SQP methods.

The aim of the present paper is to analyze the behavior of local primal-dual solutions of the log-barrier method for $s \downarrow 0$ under the Mangasarian-Fromovitz constraint qualification, some second-order sufficient optimality condition and some complementarity condition.

First we study the log-barrier method applied to the linear program constructed by a reduced local linearization of problem (1). For these auxiliary problems we obtain that the uniquely defined barrier multiplier converges to some well-defined multiplier μ of (1) associated with x^*. In fact, this is the analytic center of the multiplier set, and so our result corresponds to a similar observation in linear programming [1].

Next we show, for small s, the existence of a global minimizer $x(s)$ of $F(\cdot, s)$ on the set of strictly feasible points of some neighborhood of x^* such that $x(s)$ is a locally isolated stationary point of $F(\cdot, s)$ and defines a continuously differentiable primal trajectory. Moreover, the associated multipliers $y(s)$ converge to the limit μ obtained via the reduced linearization. This partially recovers results recently presented in [5, 21], where similar properties of the primal-dual path under MFCQ were derived. However, our approach is significantly different, and we derive some additional insights into the nature of the log-barrier method: The particular linearization gives a useful formula for the limit multiplier μ, and the Lipschitz estimate for the dual solutions given in Theorem 2 as well as the corresponding proof technique are new.

So we present a straightforward, complete and rather short convergence analysis of the primal-dual solutions. Our main tools consist in the subtle estimate of the values $g_i(x)$ on line-segments between two stationary points and of a basic fact on Lipschitzian behavior of primal-dual solutions for parametric nonlinear problems, compare estimate (33).

Our ideas might be of interest also for a bigger class of penalty-barrier methods. Finally, it turns out that parametric optimization presents us some powerful tools, see, e. g., [2, 4, 8, 9, 14, 17].

Throughout this paper, the following hypotheses are imposed.

(A1) x^* is some local minimizer of (1).

(A2) The *Mangasarian-Fromovitz constraint qualification* (MFCQ) is satisfied at x^*, i. e., the set

$$U^0 = \{u \in \mathbb{R}^n : \nabla g_i(x^*)^T u < 0 \ \forall i \in I_0\} \tag{6}$$

is not empty where $I_0 = \{i : g_i(x^*) = 0\}$.

We put $I_1 = \{1, \ldots, m\} \setminus I_0$ and suppose that the set I_0 of active constraints at x^* is not empty, too. As a consequence of (A1) and (A2) the Karush–Kuhn–Tucker (KKT) conditions

$$\nabla f(x^*) + \sum_{i=1}^{m} y_i \nabla g_i(x^*) = 0, \quad y \geq 0, \quad g(x^*) \leq 0, \quad y^{\mathrm{T}} g(x^*) = 0, \tag{7}$$

are satisfied with some nonempty and bounded set Y^* of Lagrange multipliers $y \in \mathbb{R}^m$. Note that MFCQ ensures $G^0 \neq \emptyset$. Hence, the feasible region of the log-barrier auxiliary problem (2) is nonempty. Further, we suppose throughout that

(A3) the *strict complementarity condition w.r. to* Y^* holds, i. e.,

$$\text{some } y^* \in Y^* \text{ satisfies } \ y_i^* > 0 \ \ \forall i \in I_0. \tag{8}$$

Let us introduce some further notation. The gradient and the Hessian of $F(\cdot, s)$ with respect to x are denoted by $\nabla F(x, s)$ and $\nabla^2 F(x, s)$, respectively. Further, the Landau symbols $O(\cdot)$ and $o(\cdot)$ are used in the sense that $t = O(\tau)$ means $|t| \leq c\tau$ for some constant $c > 0$ and $\tau \downarrow 0$, while $o(\cdot)$ means $o(\tau)/\tau \to 0$ as $\tau \downarrow 0$.

2. LOCALLY LINEARIZED PROBLEMS

In this section, we consider the linear program

$$\begin{aligned} &\tilde{f}(x) = \nabla f(x^*)^T (x - x^*) \to \min ! \\ \text{s. t.} \quad &\tilde{g}_i(x) = \nabla g_i(x^*)^T (x - x^*) \leq 0, \quad i \in I_0, \quad x - x^* \in \operatorname{span}\{\nabla g_i(x^*)\}_{i \in I_0}. \end{aligned} \tag{9}$$

Throughout this section,

we suppose the general assumptions (A1), (A2) and (A3).

Though condition (A3) (i. e., strict complementarity w.r. to some $y \in Y^*$) is certainly restrictive, we cannot avoid this, and it is rather natural in the context of log-barrier methods as its role for superlinear convergence and good sensitivity properties (even in linear programming) is well-known (cf. [3, 5, 10, 21]). Let

$$A = (\ldots \nabla g_i(x^*) \ldots)_{i \in I_0} \quad \text{(column-wise)} \tag{10}$$

denote the (n, m_0)-matrix which assembles the gradients $\nabla g_i(x^*)$, $i \in I_0$, of the constraints being active at x^*.

Further, let $\mathcal{R}(A)$ and $\mathcal{N}(A^\mathsf{T})$ denote the range of A and the null space of A^T, respectively. Recall that the direct sum $\mathcal{R}(A) \oplus \mathcal{N}(A^\mathsf{T})$ coincides with $\mathbb{R}^n$. Substituting $d = x - x^*$ and taking into account that span $\{\nabla g_i(x^*)\}_{i \in I_0} = \mathcal{R}(A)$, problem (9) can be equivalently given by

$$\nabla f(x^*)^T d \to \min ! \quad \text{s.t.} \quad d \in \mathcal{R}(A), \quad A^T d \leq 0. \tag{11}$$

Lemma 1. The point x^* is the unique solution of problem (9).

Proof. Obviously, x^* solves (9) uniquely if and only if

$$d \in \mathcal{R}(A), \quad A^T d \leq 0, \quad d \neq 0 \quad \Longrightarrow \quad \nabla f(x^*)^T d > 0. \tag{12}$$

Because $\mathbb{R}^n = \mathcal{R}(A) \oplus \mathcal{N}(A^T)$ we have $\mathcal{R}(A) \cap \mathcal{N}(A^T) = \{0\}$. This implies, for some $c > 0$,

$$\|A^T d\|_\infty \geq c\|d\|, \quad \forall\, d \in \mathcal{R}(A). \tag{13}$$

Thus, with the multiplier y^* of (A3), already the assertion follows

$$d \in \mathcal{R}(A), \quad A^T d \leq 0, \quad d \neq 0 \Rightarrow -\nabla f(x^*)^T d = \langle y^*, A^T d \rangle < 0. \qquad \square$$

Next, for the sake of comparison we study the log-barrier method to the reduced problem (11) (or (9), respectively), i. e. we investigate auxiliary problems

$$\begin{gathered} \varphi_s(d) = \nabla f(x^*)^T d - s \sum_{i \in I_0} \ln(-\nabla g_i(x^*)^T d) \to \min ! \\ \text{subject to} \quad d \in D^0 = \{\, d \in \mathcal{R}(A) \;:\; A^T d < 0 \,\}. \end{gathered} \tag{14}$$

By MFCQ and $\mathcal{R}(A) \oplus \mathcal{N}(A^\mathsf{T}) = \mathbb{R}^n$, D^0 is nonempty. Obviously, the objective φ_s (for any fixed $s > 0$) of (14) is strictly convex on D^0.

Lemma 2. For any $s > 0$ problem (14) possesses a unique solution $\tilde{d}(s)$. Further, there is a unique solution d^* of the problem

$$\prod_{i \in I_0} \frac{-\nabla g_i(x^*)^T d}{\nabla f(x^*)^T d} \to \max ! \quad \text{s.t.} \quad d \in D^0, \quad \|d\| = 1, \tag{15}$$

and it holds $\tilde{d}(s) = t_s d^*$ with some $t_s > 0$ for all $s > 0$ as well as $\|\tilde{d}(s)\| = O(s)$.

Proof. With (A3), we have (12) as shown in the proof of Lemma 1. Together with the growth of the logarithmic function this guarantees the existence of a solution of problem (14). The strict convexity of the objective function then implies the uniqueness of the solution.

To show the second statement, let $s > 0$ be fixed. Because of the cone-structure of D^0, any $d \in D^0$ can be represented by

$$d = t\,\hat{d} \quad \text{with some} \quad \hat{d} \in D^0, \quad \|\hat{d}\| = 1 \quad \text{and} \quad t > 0.$$

First, let us fix some arbitrary $\hat{d} \in D^0$ with $\|\hat{d}\| = 1$. Along the related ray, (14) becomes

$$\varphi_s(t\hat{d}) = t\,\nabla f(x^*)^T \hat{d} - s \sum_{i \in I_0} \ln(t\,(-\nabla g_i(x^*)^T \hat{d})) \to \min{}_t\,! \quad \text{s.t.} \quad t > 0. \tag{16}$$

This has a unique minimizer

$$\hat{t} = \hat{t}(\hat{d}) = \frac{s\,m_0}{\nabla f(x^*)^T \hat{d}}, \quad \text{where} \quad m_0 = \operatorname{card} I_0.$$

Inserting $\hat{t} = \hat{t}(\hat{d})$ into the objective of (16), this yields

$$\varphi_s(\hat{t}\hat{d}) = s\,m_0 - s \sum_{i \in I_0} \ln \left(s\,m_0 \frac{-\nabla g_i(x^*)^T \hat{d}}{\nabla f(x^*)^T \hat{d}} \right).$$

Thus, the minimum $\nu(s)$ of $\varphi_s(\hat{t}\hat{d})$ with respect to all $\hat{d} \in D^0$ satisfying $\|\hat{d}\| = 1$ (which corresponds to the minimum of (14)) is attained, and hence a solution d^* of (15) exists and realizes $\nu(s)$. By applying the above arguments to $\hat{d} = d^*$, one has that

$$\tilde{d}(s) = t_s d^*, \quad \text{with} \quad t_s = \frac{s\,m_0}{\nabla f(x^*)^T d^*}, \tag{17}$$

solves (14). Since problem (14) has a unique solution, this guarantees also the uniqueness of the solution of (15). Finally, (17) immediately implies $\|\tilde{d}(s)\| = O(s)$. □

Note that (14) is equivalent to

$$\tilde{F}(x,s) = \tilde{f}(x) - s \sum_{i \in I_0} \ln(-\tilde{g}_i(x)) \to \min\,! \quad \text{subject to} \quad x \in \{x^*\} + D^0. \tag{18}$$

With $\tilde{d}(s)$ from Lemma 2, $\tilde{x}(s) = x^* + \tilde{d}(s)$ is of course the unique solution of (18) for all $s > 0$ and satisfies $\|\tilde{x}(s) - x^*\| = O(s)$. Moreover, for $d \in \mathcal{N}(A^T)$ the KKT conditions at x^* yield $\nabla f(x^*)^T d = 0$ and hence

$$(\nabla \tilde{F}(\tilde{x}(s), s)^T d = (\nabla f(x^*) - \sum_{i \in I_0} \frac{s}{\tilde{g}_i(\tilde{x}(s))} \nabla g_i(x^*))^T d = 0.$$

On the other hand, $\nabla \tilde{F}(\tilde{x}(s), s)^T d = 0$ obviously holds for all $d \in \mathcal{R}(A)$. Therefore, $\tilde{x}(s)$ also satisfies $\nabla \tilde{F}(\tilde{x}(s), s) = 0$ for $s > 0$.

Theorem 1. The log-barrier method (14) yields for the barrier multipliers $\tilde{y}(s)$ related to the solutions $\tilde{x}(s)$ that

$$\tilde{y}_i(s) = \frac{-s}{\tilde{g}_i(\tilde{x}(s))} \equiv \mu_i, \quad i \in I_0, \tag{19}$$

where

$$\mu_i = -\frac{1}{m_0}\frac{\nabla f(x^*)^T d^*}{\nabla g_i(x^*)^T d^*}, \; i \in I_0 \quad \text{and } d^* \text{ solves (15).} \tag{20}$$

Setting $\mu_i = 0 \; \forall i \in I_1$, μ is a multiplier of the original problem (1).

Proof. From the definition of $\tilde{y}_i(s)$ and $\tilde{d}(s) = t_s d^*$ in the proof of Lemma 2, one has

$$\tilde{y}_i(s) = -\frac{s}{\tilde{g}_i(\tilde{x}(s))} = -\frac{s}{\nabla g_i(x^*)^T \tilde{d}(s)} = -\frac{s}{\nabla g_i(x^*)^T (t_s d^*)} \,.$$

Now, from $t_s = \frac{s\, m_0}{\nabla f(x^*)^T d^*}$ according to (17) the formula (19) directly follows.

To see the last assertion, we notice that without the normalization $\|d\| = 1$, problem (15) is equivalent to

$$\sum_{i\in I_0} \ln(-\nabla g_i(x^*)^T d) - m_0 \ln(\nabla f(x^*)^T d) \to \max ! \quad \text{s.t.} \quad d \in D^0.$$

The related optimality condition

$$-\sum_{i\in I_0} \frac{1}{\nabla g_i(x^*)^T d} \nabla g_i(x^*) + m_0 \frac{1}{\nabla f(x^*)^T d} \nabla f(x^*) = 0$$

yields

$$\sum_{i\in I_0} \frac{-1}{m_0} \frac{\nabla f(x^*)^T d}{\nabla g_i(x^*)^T d} \nabla g_i(x^*) + \nabla f(x^*) = 0.$$

With $d = d^* \in D^0$, this is just a specific realization of the KKT-conditions of the original problem. □

Remark 1. The multiplier μ appearing in (20) is the analytic center of the set Y^*, i.e., it is the solution of the problem

$$\Pi_{\,i\in I_0}\, y_i \to \max ! \quad \text{subject to } y \in Y^*. \tag{21}$$

In game theory, such problems are used for describing Nash-bargaining solutions over convex sets, cf. [15]. Note that (A2) and (A3) guarantee the existence of a solution $\hat{y}$ of (21) which is uniquely characterized by $\hat{y} \in Y^*$ and the Lagrange condition

$$\hat{y}_i^{-1} = -\nabla g_i(x^*)^{\mathsf{T}} \hat{\lambda} > 0 \quad \text{for some } \hat{\lambda} \text{ and all } i \in I_0. \tag{22}$$

Setting $\hat{\lambda} = td^*$ with $t = m_0(\nabla f(x^*)^{\mathsf{T}} d^*)^{-1}$ in the above theorem shows $\mu = \hat{y}$. This recovers a known result (cf., e.g., [1, 21]) by giving a concrete form and computational rule.

3. THE NONLINEAR PROBLEM

In this section, we study the local convergence of primal and dual solutions in the log-barrier method (2) near x^*,

$$F(x,s) = f(x) - s\sum_{i=1}^{m} \ln(-g_i(x)) \to \ \min !$$
$$\text{subject to} \quad x \in G^\varepsilon = \{\, x \in \mathbb{R}^n \ : \ g_i(x) < 0,\ i = 1,\ldots,m,\ \|x - x^*\| < \varepsilon\},$$

for solving the original problem (1). We assume throughout that the assumptions (A1), (A2) and (A3) are satisfied, and in addition

(A4) the following second-order optimality condition holds:

$$u^\mathsf{T} D_x^2 L(x^*,y)u > 0 \text{ for all } y \in Y^* \text{ and all } u \in U^*, u \neq 0,$$

where

$$U^* = \{u : \nabla f(x^*)^\mathsf{T} u = 0,\ \nabla g_i(x^*)^\mathsf{T} u \le 0\ \forall\ i \in I_0\}$$

is the critical cone for x^*.

Recall that $L(x,y)$ means the Lagrange function, I_0 is the active index set at x^*, and Y^* is the multiplier set associated with x^*. Note that assumption (A3) implies that U^* has the form

$$U^* \quad = \quad \{u \in \mathbb{R}^n : \nabla g_i(x^*)^\mathsf{T} u = 0\ \forall\ i \in I_0\}\,. \tag{23}$$

The stationary points $x(s)$ of $F(\cdot, s)$ on G^ε are the zeros of

$$\nabla F(x,s) = D_x L(x,y) \ \text{ with } y_i = -s/g_i(x) > 0 \quad \forall i. \tag{24}$$

To see that the related log-barrier multipliers $y(s)$ are bounded for small s and ε, consider any $u \in U^0$ (6). If ε was small enough, we have $\nabla g_i(x(s))^\mathsf{T} u < -r_u < 0$ $\forall i \in I_0$ and

$$\nabla F(x(s),s)^\mathsf{T} u = \nabla f(x(s))^\mathsf{T} u + \sum_{i\in I_0} y_i(s)\nabla g_i(x(s))^\mathsf{T} u + \sum_{k\in I_1} y_k(s)\nabla g_k(x(s))^\mathsf{T} u = 0.$$

Since also

$$y_k(s) \le 2s \Big/ \left|\max_{k\in I_1}\ g_k(x^*)\right| \to 0 \tag{25}$$

is valid for $k \in I_1$ and small ε (if $I_1 \neq \emptyset$), one obtains that

$$y_i(s) = -s/g_i(x(s)) < C_g \quad \text{ and } \quad g_i(x(s)) < -s/C_g \quad \forall i \tag{26}$$

hold with some constant $C_g > 0$ and for sufficiently small s, ε. Further (26) and $I_0 \neq \emptyset$ imply (since g is locally Lipschitz)

$$\|x(s) - x^*\| \ge Cs \quad \text{with some constant } C. \tag{27}$$

So the convergence $x(s) \to x^*$, if valid at all, is not very fast. In what follows we will show

Theorem 2. There are $\bar{s} > 0$ and $\varepsilon > 0$ such that for all $s \in (0, \bar{s})$, the function $F(\cdot, s)$ on G^ε has a global minimizer $x(s)$ which is the unique stationary point of $F(\cdot, s)$ on G^ε. The associated multipliers $y(s)$ (5) converge to μ given in Theorem 1 where

$$\mathrm{dist}\,((x(s), y(s)), (x^*, Y^*)) \;\leq\; C^* s \text{ with some constant } C^*, \tag{28}$$

the Hessian $\nabla^2 F(x(s), s)$ is uniformly positive definite and $x(\cdot)$ is continuously differentiable on $(0, \bar{s})$. Theorem 2 will be proved via several lemmata. Our main tools consist of elementary (but quite sharp) estimates of the values $g_i(x)$ on line-segments between two stationary points $x(s)$ and $\xi(s)$ and of an basic fact on Lipschitzian solutions for parametric problems, cf. (33). Notice that different approaches of dealing with barrier (and penalty) methods in the framework of parametric optimization (but under different hypotheses) can be found, e. g., in [8, 11, 12, 13].

Lemma 3. There is some $\varepsilon > 0$ such that, for sufficiently small s, the following holds:

(i) The function $F(\cdot, s)$ has a global minimizer on G^ε, and each stationary solution $x(s) \in G^\varepsilon$ of $F(\cdot, s)$ and its associated multiplier $y(s)$ according to (5) satisfy a Lipschitz estimate

$$\|x(s) - x^*\| \;=\; O(s) \quad \text{and} \quad \mathrm{dist}\,(y(s), Y^*) \;=\; O(s). \tag{29}$$

(ii) Moreover, there are positive constants K, K_1, K_2 such that, for any (possibly second) stationary solution $\xi(s)$, the points x of the connecting line-segment $[x(s), \xi(s)]$ belong to G^ε and satisfy

$$\|x - x^*\| \leq Ks \quad \text{and} \quad -K_1 s \leq g_i(x) \leq -K_2 s \quad \forall\, i \in I_0. \tag{30}$$

P r o o f. First we derive the existence result and $\lim \mathrm{dist}\,((x(s), y(s)), x^* \times Y^*) = 0$. Then we prove (29), and finally the estimates (30) are shown. For brevity, we write in this proof x^s, y^s and ξ^s instead of $x(s)$, $y(s)$ and $\xi(s)$, respectively.

Part 1, (i). Under (A2) and (A4) the point x^* is an isolated local minimizer (even an isolated stationary solution) of the original problem (1), see [17]. Having this, the proof of the existence of a global minimizer x^s of $F(\cdot,s)$ on G^ε such that $x^s \to x^*$ as $s \downarrow 0$ is standard, see, e. g., Fiacco and McCormick [4, Thm. 10], and will be omitted.

By (26) we know that all $y_i^s = -s/g_i(x^s)$ are bounded. Thus, accumulation points (x^0, y^0) of arbitrary stationary pairs (x^s, y^s) as $s \downarrow 0$ exist and are KKT-points (even if $g_i(x^0) < 0$ for some $i \in I_0$) for the initial problem (1). As mentioned above, x^* is an isolated stationary point for (1), cf. [17]. Hence, for small ε, we have $x^0 = x^* = \lim x^s$ and, in consequence, also $\lim \mathrm{dist}\,(y^s, Y^*) = 0$.

Part 2, (i). Now we estimate the quantities $\|x^s - x^*\|$ and $\operatorname{dist}(y^s, Y^*)$ in terms of s and $g(x^s)$: Given (x^s, y^s), put

$$a(s) = \sum_{i \notin I_0} y_i^s \nabla g_i(x^s), \quad b_i(s) = \begin{cases} g_i(x^s), & \text{if } i \in I_0 \\ 0, & \text{if } i \notin I_0. \end{cases} \tag{31}$$

Then $\nabla f(x^s) + \sum_{i \in I_0} y_i^s \nabla g_i(x^s) + a(s) = 0$ yields that $(x^s, \bar{y}(s))$ with

$$\bar{y}_i(s) = y_i^s \;\text{ if } i \in I_0, \quad \bar{y}_k(s) = 0 \;\text{ if } k \in I_1,$$

is a KKT point for problem

$$f(x) - a(s)^{\mathsf{T}} x \to \min ! \quad \text{s.t.} \quad g(x) \le b(s). \tag{32}$$

By (25) we have $\|a(s)\| \le s\, C_a$ with some constant C_a and, as just shown for the crucial components $i \in I_0$, $\lim b(s) = 0$. By Corollary 2.9 and Theorem 8.36 in [11] or by the Theorems 2.2, 2.4, 3.1, 3.2 and 4.2 in Robinson's basic paper [17], this ensures, for small s, that x^s is a local minimizer of (32) and that, since $\|(a, b)\| \to 0$ as $s \downarrow 0$, some Lipschitz estimate

$$\operatorname{dist}((x^s, \bar{y}(s)), (x^*, Y^*)) \;\le\; C_{KKT} \|(a(s), b(s))\| \tag{33}$$

holds true for small s and some constant C_{KKT}. Recalling (25) the estimate (33) is also valid (possibly with a new constant) for (x^s, y^s).

In this moment, we do not know whether also $b(s)$ from (31) satisfies a Lipschitz estimate. In fact, this statement is more involved and needs the extra assumption (A3).

Put $\tau = \|x^s - x^*\|$, and let y^* be the multiplier in (A3); then $y_i^* > 0 \; \forall i \in I_0$. Since $D_x L(x^*, y^*) = 0$, we obtain

$$L(x^s, y^*) - L(x^*, y^*) = o(\tau)$$

Because of $y^{*\mathsf{T}} g(x^*) = 0$, this means

$$f(x^s) - f(x^*) + \sum_i y_i^* g_i(x^s) = o(\tau). \tag{34}$$

Similarly, $\operatorname{dist}(y^s, Y^*) \to 0$ ensures $L(x^s, y^s) - L(x^*, y^s) = o(\tau)$, i.e.,

$$f(x^s) - f(x^*) + \sum_{i=1}^{m} y_i^s g_i(x^s) - \sum_{i=1}^{m} y_i^s g_i(x^*) = o(\tau).$$

Since $\sum_{i=1}^{m} y_i^s g_i(x^s) = -ms$ (compare (5)) and $g_i(x^*) = 0 \; \forall i \in I_0$, the latter is

$$f(x^s) - f(x^*) = ms + \sum_{k \in I_1} y_k^s g_k(x^*) + o(\tau). \tag{35}$$

Comparing (34) and (35) and using $\sum_{k\in I_1} y_k^s g_k(x^*) = O(s)$, we finally obtain

$$-\sum_{i\in I_0} y_i^* g_i(x^s) = O(s) + o(\tau). \tag{36}$$

Since all y_i^*, $i \in I_0$, are positive, the sum $-\sum_{i\in I_0} y_i^* g_i(x^s)$ can be taken as the norm of $b(s)$, defined in (31). Now (33) yields

$$\operatorname{dist}((x^s, y^s), (x^*, Y^*)) \leq C_{KKT}\|(a(s), b(s))\| = O(s) + o(\tau) \tag{37}$$

and, in particular, $\tau = O(s) + o(\tau)$. For small s such that $\|o(\tau)\| < \frac{1}{2}\tau$ we thus conclude $\tau < 2O(s) = O(s)$, and since g is locally Lipschitz, $g_i(x^s) = O(s)\ \forall\, i \in I_0$, too. Therefore, under (A3), the estimate (33) holds in the form (28)

$$\operatorname{dist}((x^s, y^s), (x^*, Y^*)) \leq C^* s \quad (s \text{ small}).$$

Taking (26) into account, we thus obtain (30) for sufficiently small s and related stationary points x^s.

Part 3, (ii). Now let $x \in [x^s, \xi^s]$ where x^s, ξ^s are stationary solutions for small s. To show $x \in G^\varepsilon$, it suffices to study all g_i, $i \in I_0$, and to verify (30) for the related points x with possibly new constants. Clearly, since x^s and ξ^s fulfill (30), the first estimate holds for x with constant $K' = K$, too. Using a local Lipschitz rank L of g near x^*, we may put $K_1' = LK$ due to $-LKs \leq g_i(x)$. For verifying the third inequality, put $t_s = \|\xi^s - x^s\|$, $u^s = (\xi^s - x^s)/t_s$, $x = x^s + tu^s$ and suppose $g_i(x^s) \leq g_i(\xi^s)$ (without loss of generality). We have to deal with the case of

$$g_i(x) = g_i(x^s + tu^s) \notin J(s) = [g_i(x^s), g_i(\xi^s)] \text{ for some } i \in I_0 \text{ and } t \in (0, t_s), \tag{38}$$

otherwise nothing remains to prove. Hence assume that

$$\max_{0\leq t\leq t_s} g_i(x^s + tu^s) \notin J(s) \quad \text{or} \quad \min_{0\leq t\leq t_s} g_i(x^s + tu^s) \notin J(s)$$

We consider the first case, the other one can be handled analogously. Since any maximizer t^* fulfills $0 < t^* < t_s$, we obtain $\nabla g_i(x^s + t^*u^s)^\mathsf{T} u^s = 0$. So it follows with some local Lipschitz rank L' of Dg near x^*,

$$|\nabla g_i(x^s + \theta u^s)^\mathsf{T} u^s| \leq L' t_s \quad \forall\, \theta \in (0, t_s).$$

By the mean value theorem, some $\theta \in (0, t_s)$ satisfies

$$|g_i(x^s + t^*u^s) - g_i(x^s)| = t^*|\nabla g_i(x^s + \theta u^s)^\mathsf{T} u^s| \leq t^* L' t_s \leq L' t_s^2. \tag{39}$$

Since $t_s \leq \|\xi^s - x^*\| + \|x^s - x^*\| \leq 2Ks$, the latter implies

$$|g_i(x) - g_i(x^s)| \leq 4L'K^2 s^2 \tag{40}$$

and guarantees (30) with the third new constant $K_2' = \frac{1}{2}K_2$. □

Our proof in Part 3 has been made in such a way that the following conclusion becomes evident: If $x \in [x(s), \xi(s)]$ fulfills (38) then (40) holds true. This yields the next lemma as a direct application of (30).

Lemma 4. Let $z_i(x,s) = s/g_i(x)$ for $x \in G^\varepsilon, i \in I_0$. If, under the assumptions of Lemma 3, there is a common limit $z_i^* = lim_{s\downarrow 0}\, z_i(x,s)$ for the two settings $x = x(s)$ and $x = \xi(s)$, then the limit exists and remains the same for all $x \in [x(s), \xi(s)]$.

Lemma 5. For sufficiently small $s \downarrow 0$, let x^s belong to a line-segment $[x(s), \xi(s)]$ connecting stationary points of $F(\cdot, s)$ on G^ε as in Lemma 3. Then the multipliers y^s with $y_i^s = -s/g_i(x^s)\ \forall i$ converge to μ given by Theorem 1.

Proof. We suppose first that $x^s = x(s)$ is stationary. For any sequence of certain $s = s_k \downarrow 0$, let η be some accumulation point of the related duals y^s which are bounded due to (25) and (30) and fulfill $y_k^s \to 0$ for $k \in I_1$. Let $i \in I_0$ and put

$$\tau_s = \|x^s - x^*\| \quad \text{and} \quad \lambda^s = (x^s - x^*)/\tau_s.$$

Without loss of generality let $y^s \to \eta$ and $\lambda^s \to \lambda$ already hold for the initial sequence of $s = s_k$. Then $g_i(x^*) = 0$ and $\lambda^s \to \lambda$ yield with (30)

$$g_i(x^s) = \tau_s \nabla g_i(x^*)^{\mathsf T} \lambda^s + o(\tau_s) \text{ and } \tau_s^{-1} g_i(x^s) \to \nabla g_i(x^*)^{\mathsf T}\lambda < 0.$$

Further (taking a subsequence if necessary), τ_s/s has, by (30) and (27), a limit $\gamma > 0$. So it follows

$$\frac{g_i(x^s)}{s} = \frac{g_i(x^s)}{\tau_s}\,\frac{\tau_s}{s} \to \gamma\, \nabla g_i(x^*)^{\mathsf T}\lambda, \tag{41}$$

$$\eta_i^{-1} = -\gamma\, \nabla g_i(x^*)^{\mathsf T}\lambda > 0. \tag{42}$$

In addition, $\eta \in Y^*$ follows from stationarity $D_x L(x^s, y^s) = 0$. Therefore, Remark 1 ensures $\eta = \mu$. Since (42) was obtained for any sequence of $s = s_k \downarrow 0$, and the analytic center is unique, the statement of the present lemma holds for $x^s = x(s)$ and $x^s = \xi(s)$. Taking Lemma 4 into account, the proof is complete. □

Completing the Proof of Theorem 2:

Because of the preceding lemmata, it remains to show that $x(s)$ is the unique stationary solution of $F(\cdot, s)$ on G^ε, and, moreover, the Hessian $\nabla^2 F(x(s), s)$ is uniformly positive definite and $x(\cdot)$ is continuously differentiable on $(0, \bar s)$.

Let G^ε be according to Lemma 3 and let all derivatives and function values be taken at $x \in G^\varepsilon$ and $z = z(x,s)$ with $z_i(x,s) = -s/g_i(x)\ \forall i$. The Hessian $\nabla^2 F(x,s)$ (w.r. to x) becomes

$$\begin{aligned} \nabla^2 F(x,s) &= \nabla^2 f(x) - \textstyle\sum_i \frac{s}{g_i(x)} \nabla^2 g_i(x) + \sum_i \frac{s}{g_i^2(x)} \nabla g_i(x) \nabla g_i(x)^{\mathsf T} \\ &= D_x^2 L(x,z) + \frac{1}{s} \textstyle\sum_i (z_i \nabla g_i(x))(z_i \nabla g_i(x)^{\mathsf T}). \end{aligned} \tag{43}$$

Next we ask for uniform positive definiteness of $\nabla^2 F(x,s)$ for small s and ε where x belongs to some line-segment $x \in [x(s), \xi(s)]$ connecting two stationary points $x(s)$

and $\xi(s)$ of $F(\cdot, s)$ on G^ε. The crucial matrix $C(x,z) = \sum_i (z_i \nabla g_i)(z_i \nabla g_i)^\mathsf{T}$ is a (positive semidefinite) dyadic product. Row p consists of the elements

$$c_{p,q} = \sum_i \frac{z_i \partial g_i(x)}{\partial x_p} \frac{z_i \partial g_i(x)}{\partial x_q}.$$

We are going to analyze regularity and limits of $C(x,z)$ now. For $s \downarrow 0$, we have $x \to x^*$ according to Lemma 3 while Lemma 5 ensures that all $z(x,s)$ converge to $z^* = \mu$ in Theorem 1. The related matrices $C(x,z)$ then converge to

$$C^* = \sum_i (z_i^* A_i)(z_i^* A_i^\mathsf{T}) \quad \text{where } A_i = \nabla g_i(x^*).$$

This matrix fulfills

$$\ker C^* = \bigcap_{i:\ z_i^* > 0} \ker A_i. \tag{44}$$

Indeed, if $w \in \ker C^*$ then $0 = w^\mathsf{T} C^* w = \sum_i (z_i^*)^2 \langle A_i, w\rangle^2$. Therefore, $z_i^* \neq 0$ implies $A_i w = 0$. Conversely, if $w \in \cap_{i:\ z_i^*>0} \ker A_i$ then $C^* w = 0$ holds trivially. This proves (44).

Since $z_i^* > 0$ iff $i \in I_0$, the kernel of C^* is as small as possible

$$\ker C^* = \cap_{i \in I_0} \ker A_i = U^*. \tag{45}$$

Next we apply continuity of $D_x^2 L$ and $C(x,z)$ at (x^*, z^*).

Provided that ε and s are small enough, the second order condition (A4) and (45) ensure that there are positive β, γ such that $\operatorname{dist}(u, U^*) < \beta$ and $\|u\| = 1$ yield $u^\mathsf{T} D_x^2 L(x,z) u \geq \gamma > 0$ for all $x \in G^\varepsilon \cup [x(s), \xi(s)]$ and $z = z(x,s)$. Further, $u^\mathsf{T} C(x,z) u \geq 0$ is always true. Hence we obtain

$$u^\mathsf{T} \nabla^2 F(x,s) u = u^\mathsf{T} D_x^2 L(x,z) u \ + \frac{1}{s} u^\mathsf{T} C(x,z) u \geq \gamma.$$

For the remaining normalized u (with $\operatorname{dist}(u, U^*) \geq \beta$) and the same (x,z), it holds both $u^\mathsf{T} C(x,z) u \geq \gamma'$ with some $\gamma' > 0$, and $u^\mathsf{T} D_x^2 L(x,z) u \geq q$ with some fixed q. This ensures

$$u^\mathsf{T} \nabla^2 F(x,s) u = u^\mathsf{T} D_x^2 L(x,z) u \ + \frac{1}{s} u^\mathsf{T} C(x,z) u \geq q + \frac{\gamma'}{s}.$$

Since these constants do not depend of s, the matrix

$$\nabla^2 F(x,s) \text{ is uniformly positive definite for small } s \text{ and } x \in [x(s), \xi(s)].$$

Notice that Lemma 3 guarantees $[x(s), \xi(s)] \subset G^\varepsilon$. So, if $x(s) \neq \xi(s)$, we may use that ∇F is continuously differentiable on the the segment. Writing $\xi(s) = x(s) + t(s) u(s)$, $\|u(s)\| = 1, t(s) \downarrow 0$ one finally obtains $\nabla F(\xi(s), s) = \nabla F(x(s), s) = 0$ as well as

$$t(s) \nabla^2 F(x(s), s) u(s) \ = \ o(t(s)).$$

For $s \downarrow 0$, so $\nabla^2 F(x^s, s)u(s)$ vanishes, a contradiction to uniform definiteness. This tells us that $x(s) = \xi(s)$.

Evidently, the uniform positive definiteness particularly says that the assumptions of the implicit function theorem for the system of equations $\nabla F(x, s) = 0$ are satisfied at each $x = x(s)$ if $s \in (0, \bar{s})$. Hence $x(\cdot)$ is continuously differentiable on $(0, \bar{s})$. □

As shown in the literature (cf. e.g. [16, 19]) primal-dual interior point methods based on log-barriers approximate the primal-dual path. Hence, the results derived here for the primal-dual path are also applicable to these methods and provide a convergence analysis of the generated log-barrier multipliers.

(Received November 14, 2003.)

REFERENCES

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

[2] B. Bank, J. Guddat, D. Klatte, B. Kummer, and K. Tammer: Non-linear Parametric Optimization. Akademie–Verlag, Berlin 1982.

[3] A. S. El-Bakry, R. A. Tapia, and Y. Zhang: On the convergence rate of Newton interior-point methods in the absence of strict complementarity. Comput. Optim. Appl. *6* (1996), 157–167.

[4] A. V. Fiacco and G. P. McCormick: Nonlinear Programming: Sequential Unconstrained Minimization Techniques. Wiley, New York 1968.

[5] A. Forsgren, P. E. Gill, and M. H. Wright: Interior methods for nonlinear optimization. SIAM Rev. *44* (2002), 525–597.

[6] C. Grossmann: Convergence analysis for penalty/barrier path-following in linearly constrained convex programming. Discuss. Math.: Differential Inclusions, Control and Optimization *20* (2000), 7–26.

[7] C. Grossmann and A. A. Kaplan: Strafmethoden und modifizierte Lagrange Funktionen in der nichtlinearen Optimierung. Teubner, Leipzig 1979.

[8] C. Grossmann and G. Schöniger: Sensitivität und Anwendbarkeit von Straf–Barriere–Methoden. Z. Angew. Math. Mech. *57* (1977), 419–430.

[9] J. Guddat, F. Guerra, and D. Nowack: On the role of the Mangasarian-Fromovitz constraint qualification for penalty-, exact penalty-, and Lagrange multiplier methods. In: Mathematical programming with data perturbations (A. Fiacco, ed.), (Lecture Notes in Pure and Appl. Mathematics 195). Marcel Dekker, Dordrecht 1997, pp. 159–183.

[10] O. Güler: Limiting behavior of weighted central paths in linear programming. Math. Programming *65A* (1994), 347–363.

[11] D. Klatte and B. Kummer: Nonsmooth Equations in Optimization – Regularity, Calculus, Methods and Applications. Kluwer, Dordrecht 2002.

[12] B. Kummer: On solvability and regularity of a parametrized version of optimality conditions. Z. Oper. Res.: Math. Meth. Oper. Res. *45* (1995), 215–230.

[13] B. Kummer: Parametrizations of Kojima's system and relations to penalty and barrier functions. Math. Programming, Ser. B *76* (1997), 579–592.

[14] F. Nožička, J. Guddat, H. Hollatz, and B. Bank: Theorie der linearen parametrischen Optimierung. Akademie–Verlag, Berlin 1974.

[15] G. Owen: Spieltheorie. Springer–Verlag, Berlin 1971 (Translation).

[16] D. Ralph and S. J. Wright: Superlinear convergence of an interior point method despite dependent constraints. Math. Oper. Res. *25* (2000), 179–194.

[17] S. M. Robinson: Generalized equations and their solutions. Part II: Applications to nonlinear programming. Math. Programming Stud. *19* (1982), 200–221.
[18] S. J. Wright: Superlinear convergence of a stabilized SQP method to a degenerate solution. Comput. Optim. Appl. *11* (1998), 253–275.
[19] S. J. Wright: On the convergence of the Newton/log-barrier method. Math. Programming *90A* (2001), 71–100.
[20] S. J. Wright: Effects of finite-precision arithmetic on interior-point methods for nonlinear programming. SIAM J. Optim. *12* (2001), 36–78.
[21] S. J. Wright and D. Orban: Properties of the Log-barrier Function on Degenerate Nonlinear Programs. Preprint ANL/MCS-P772-0799, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne 1999, revised May 2001.

Christian Grossmann, Institut für Numerische Mathematik, Technische Universität Dresden, D-01062 Dresden. Germany.
e-mail: grossm@math.tu-dresden.de

Diethard Klatte, Institut für Operations Research, Universität Zürich, Moussonstrasse 15, CH-8044, Zürich. Switzerland.
e-mail: klatte@ior.unizh.ch

Bernd Kummer, Institut für Mathematik, Humboldt-Universität zu Berlin, Math.-Nat.-Fak. II, D-10099 Berlin. Germany.
e-mail: kummer@mathematik.hu-berlin.de