
A Note on Error Estimates for some Interior Penalty Methods ^{*}

A. F. Izmailov¹ and M. V. Solodov²

¹ Moscow State University, Faculty of Computational Mathematics and Cybernetics, Department of Operations Research, Leninskiye Gori, GSP-2, 119992 Moscow, Russia. izmaf@ccas.ru

² Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, RJ 22460-320, Brazil. solodov@impa.br

Summary. We consider the interior penalty methods based on the logarithmic and inverse barriers. Under the Mangasarian-Fromovitz constraint qualification and appropriate growth conditions on the objective function, we derive computable estimates for the distance from the subproblem solution to the solution of the original problem. Some of those estimates are shown to be sharp.

1 Introduction and Preliminaries

We consider the optimization problem

$$\begin{aligned} & \text{minimize } f(x) \\ & \text{subject to } x \in D = \{x \in \mathbb{R}^n \mid G(x) \leq 0\}, \end{aligned} \tag{1}$$

where the set

$$D^0 = \{x \in \mathbb{R}^n \mid G(x) < 0\}$$

is assumed to be nonempty. Under the stated assumption, one of the classical schemes [3] for solving problem (1) is the *interior penalty (or barrier) method*. It consists in replacing (1) by a sequence of (in some sense, unconstrained) subproblems of the form

$$\begin{aligned} & \text{minimize } \varphi_\sigma(x) \\ & \text{subject to } x \in D^0, \end{aligned} \tag{2}$$

where $\sigma > 0$ is the penalty (barrier) parameter, and

^{*} Research of the first author is supported by Russian Foundation for Basic Research Grant 04-01-00341 and RF President's Grant NS-1815.2003.1 for the support of leading scientific schools. The second author is supported in part by CNPq Grants 300734/95-6 and 471780/2003-0, by PRONEX–Optimization, and by FAPERJ.

$$\varphi_\sigma : D^0 \rightarrow \mathbb{R}, \quad \varphi_\sigma(x) = f(x) + \sigma \sum_{i=1}^m b(-G_i(x)),$$

with $b : \mathbb{R}_{++} \rightarrow \mathbb{R}_{++}$ being the barrier function. The most popular are the logarithmic barrier

$$b(t) = -\ln t, \quad (3)$$

and the inverse barrier

$$b(t) = 1/t. \quad (4)$$

For basic convergence results of this type of methods, we refer the reader to [3, 10, 11]. Here, we only mention that to ensure convergence, the barrier parameter σ must be driven to zero.

For each $\sigma > 0$, let x_σ be a solution of (2). Let \bar{x} be a solution of (1), and suppose that $x_\sigma \rightarrow \bar{x}$ as $\sigma \rightarrow 0+$. We note that our analysis can easily treat the case when there is a given sequence $\{\sigma_k\} \subset \mathbb{R}_+$ such that $\sigma_k \rightarrow 0+$, and the corresponding sequence $\{x^k\} = \{x_{\sigma_k}\}$ is convergent to \bar{x} . The modifications to cover this case are straightforward. We assume that the objective function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and the constraint mapping $G : \mathbb{R}^n \rightarrow \mathbb{R}^m$ are differentiable, and the derivatives of f and G are continuous at \bar{x} . For some results, the problem data will further be assumed twice differentiable.

In this paper, we are interested in estimates of the distance from x_σ to \bar{x} via some computable quantity, that is, in *error bounds* of the form

$$\|x_\sigma - \bar{x}\| = O(r(x_\sigma, \sigma)), \quad (5)$$

where $r : \mathbb{R}^n \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is some (easily) computable function such that, at least, $r(x_\sigma, \sigma) \rightarrow 0$ as $\sigma \rightarrow 0+$ and $x_\sigma \rightarrow \bar{x}$. Moreover, it is desirable that the bound (5) should be sharp, i.e., not improvable under the given assumptions. In some cases, it is possible to eliminate the dependence on x_σ in the right-hand side of (5), and then (5) can be considered as a convergence rate estimate. But in any case, computable error bounds are very useful. In particular, they provide reliable stopping tests for the related algorithms.

Denote the set of Lagrange multipliers associated with \bar{x} by

$$\mathcal{M} = \mathcal{M}(\bar{x}) = \left\{ \mu \in \mathbb{R}^m \mid \frac{\partial L}{\partial x}(\bar{x}, \mu) = 0, \mu \geq 0, \langle \mu, G(\bar{x}) \rangle = 0 \right\},$$

where

$$L(x, \mu) = f(x) + \langle \mu, G(x) \rangle, \quad x \in \mathbb{R}^n, \mu \in \mathbb{R}^m,$$

is the Lagrangian of problem (1). Recall that the *linear independence constraint qualification* (LICQ) for problem (1) at \bar{x} consists of saying that $G'_i(\bar{x})$, $i \in A$, are linearly independent, where $A = A(\bar{x}) = \{i = 1, \dots, m \mid G_i(\bar{x}) = 0\}$ is the set of constraints active at \bar{x} . Under LICQ, \mathcal{M} is necessarily a singleton. The weaker *Mangasarian–Fromovitz constraint qualification* (MFCQ) for problem (1) at \bar{x} consists of saying that there exists $\bar{\xi} \in \mathbb{R}^n$ such

that $G'_A(\bar{x})\bar{\xi} < 0$. If this condition holds, then \mathcal{M} is necessarily a nonempty polyhedral and compact set. We say that *strict complementarity* holds at \bar{x} if there exists some $\bar{\mu} \in \mathcal{M}$ such that $\bar{\mu}_A > 0$.

Our goal is to obtain computable error bounds under assumptions that do not use strict complementarity and do not invoke any CQ-type conditions stronger than MFCQ (so that, in particular, \mathcal{M} need not be a singleton).

We note that under appropriate assumptions, it can be possible to estimate the distance from an *arbitrary* $x \in \mathbb{R}^n$ to \bar{x} , independently of any specific algorithmic framework, i.e., regardless of how x was produced or chosen (sometimes these estimates are of a primal-dual nature; see below). We refer the reader to [12] for a survey of algorithm-independent error bounds and their applications. On the other hand, algorithm-based error bounds can sometimes be established under weaker or different assumptions than their algorithm-independent counterparts (e.g., [7], available at http://www.preprint.impa.br/Shadows/SERIE_A/2004/303.html).

As an algorithm-independent error bound relevant in our context, we mention the following result, based on [6, Lemma 2] and [4, Theorem 2]. Suppose that with some $\bar{\mu} \in \mathcal{M}$, the following second-order sufficient optimality condition holds:

$$\frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\mu})[\xi, \xi] > 0 \quad \forall \xi \in C \setminus \{0\},$$

where

$$C = C(\bar{x}) = \{\xi \in \mathbb{R}^n \mid G'_A(\bar{x})\xi \leq 0, \langle f'(\bar{x}), \xi \rangle \leq 0\} \tag{6}$$

is the critical cone of problem (1) at \bar{x} . Then for $(x, \mu) \in \mathbb{R}^n \times \mathbb{R}^m$ close enough to $(\bar{x}, \bar{\mu})$, it holds that

$$\|x - \bar{x}\| = O(r(x, \mu)),$$

where

$$r(x, \mu) = \left\| \left(\frac{\partial L}{\partial x}(x, \mu), \min\{\mu, -G(x)\} \right) \right\|,$$

with the minimum taken componentwise. Note that the quantity $r(x, \mu)$ is the *natural residual* of the *Karush-Kuhn-Tucker system*

$$\frac{\partial L}{\partial x}(x, \mu) = 0, \quad \mu \geq 0, \quad G(x) \leq 0, \quad \langle \mu, G(x) \rangle = 0,$$

which characterizes stationary points of problem (1) and the associated multipliers. This result does not rely on any CQ. Note, however, that if we are in some algorithmic framework, then for this result to be applicable, the given primal-dual sequence generated by the method has to converge to the specific $(\bar{x}, \bar{\mu})$ which satisfies the above condition. In our results, no assumptions about convergence of the dual part of the sequence are necessary, as long as the needed growth conditions (related to sufficient optimality conditions, see below) are satisfied. We refer the reader to [8] for other algorithm-independent

error bounds under assumptions which subsume some CQ-type conditions, as well as for a detailed discussion and comparisons of error bounds and regularity conditions for KKT systems.

The results presented in this paper are strongly inspired by [14] and [5], where the log-barrier method has been analyzed. The analysis in these works is based on the following assumptions: MFCQ, the strict complementarity condition, and the second-order sufficient optimality condition in the following strong form:

$$\frac{\partial^2 L}{\partial x^2}(\bar{x}, \mu)[\xi, \xi] > 0 \quad \forall \mu \in \mathcal{M}, \forall \xi \in C \setminus \{0\}. \quad (7)$$

In [14], the case of violation of strict complementarity (but with the other two assumptions satisfied) is discussed as well. The assertions to be stated below are weaker than those in [14] and [5], but our assumptions are different. We also assume MFCQ, but never strict complementarity. When we assume second-order sufficiency (actually, we assume a certain quadratic growth condition, but in the given setting it is equivalent to second-order sufficiency), it is in a form significantly weaker than (7), see the discussion below. On the other hand, for some results we assume convexity of the objective function and/or of the constraints.

Our analysis relies on the so-called growth conditions, which we discuss next. We say that the *linear growth* condition is satisfied at \bar{x} if there exist $\gamma > 0$ and a neighborhood U of \bar{x} such that

$$f(x) \geq f(\bar{x}) + \gamma \|x - \bar{x}\| \quad \forall x \in D \cap U. \quad (8)$$

As is well known (see, e.g., [2, Lemma 3.24]), the linear growth is guaranteed by the *first-order sufficient condition* (FOSC)

$$C = \{0\}. \quad (9)$$

Moreover, the two conditions are equivalent provided MFCQ holds at \bar{x} .

We say that the *quadratic growth* condition is satisfied at \bar{x} if there exist $\gamma > 0$ and a neighborhood U of \bar{x} such that

$$f(x) \geq f(\bar{x}) + \gamma \|x - \bar{x}\|^2 \quad \forall x \in D \cap U. \quad (10)$$

Obviously, quadratic growth is a weaker property than linear growth. If $\mathcal{M} \neq \emptyset$, then the following *second-order sufficient condition* (SOSC) becomes relevant:

$$\forall \xi \in C \setminus \{0\} \exists \mu \in \mathcal{M} \text{ such that } \frac{\partial^2 L}{\partial x^2}(\bar{x}, \mu)[\xi, \xi] > 0. \quad (11)$$

According to [2, Theorem 3.70], the latter condition is sufficient for the quadratic growth, and equivalent to it if MFCQ holds at \bar{x} .

We emphasize that if \mathcal{M} is not a singleton, the second order sufficient optimality condition (11) is significantly weaker than (7).

The approach we use in this paper is quite similar to the one employed in sensitivity theory for deriving Lipschitz and Hölder stability of optimal solutions (see [2]). However, the context of barrier methods possesses a special feature which can (and should) be taken into account: the perturbed solutions remain feasible for the original problem.

We start with considering the case of log-barrier in Section 2. For this barrier, it is possible to obtain convergence rate estimates where the right-hand side in (5) does not depend on x_σ . This is not the case for the inverse barrier, considered in Section 3, where the right-hand side in (5) involves x_σ . Nevertheless, it still gives a computable estimate.

2 Error Estimates for the Log-Barrier Method

Throughout this section, φ_σ is defined with the logarithmic barrier (3).

For each $\sigma > 0$, denote

$$\mu_\sigma = -\sigma(1/G_1(x_\sigma), \dots, 1/G_m(x_\sigma)) > 0. \tag{12}$$

By direct computation,

$$\langle \mu_\sigma, G(x_\sigma) \rangle = -m\sigma, \tag{13}$$

and by the first-order necessary optimality conditions for problem (2) at x_σ ,

$$\frac{\partial L}{\partial x}(x_\sigma, \mu_\sigma) = \varphi'_\sigma(x_\sigma) = 0. \tag{14}$$

We start with the case when (1) is a convex minimization problem. The following result is well-known (as we were informed by a referee, it probably first appeared in [1]). We include its short proof, for the sake of completeness.

Proposition 1. *Let f and G_i , $i = 1, \dots, m$, be convex. For $\sigma > 0$, let x_σ be a solution of problem (2) with the barrier function defined in (3). Then it holds that*

$$f(x_\sigma) \leq \inf_{x \in D} f(x) + m\sigma. \tag{15}$$

Proof. Associated to (1) is its *Wolfe dual* problem (e.g., see [9])

$$\begin{aligned} & \text{maximize } L(x, \mu) \\ & \text{subject to } (x, \mu) \in \Delta = \{(x, \mu) \in \mathbb{R}^n \times \mathbb{R}^m \mid \frac{\partial L}{\partial x}(x, \mu) = 0, \mu \geq 0\}. \end{aligned}$$

By weak duality [9, Theorem 8.1.3], it holds that

$$\sup_{(x, \mu) \in \Delta} L(x, \mu) \leq \inf_{x \in D} f(x). \tag{16}$$

Observe that, by (12) and (14), $(x_\sigma, \mu_\sigma) \in \Delta$. By (13), we obtain that

$$L(x_\sigma, \mu_\sigma) = f(x_\sigma) - m\sigma.$$

The assertion now follows from (16). \square

Note that in the above, solvability of problem (1) is not needed. But in our setting, estimate (15) gives

$$f(x_\sigma) \leq f(\bar{x}) + O(\sigma). \quad (17)$$

By (17), using the feasibility of x_σ in the original problem (1) and relation (8) (in the case of linear growth) or (10) (in the case of quadratic growth), we can immediately obtain convergence rate estimates, stated in Theorem 1 below. We note that estimate (19) of this theorem was obtained in [10, 11] assuming strong convexity of the Lagrangian for some fixed multiplier, which is stronger than the quadratic growth condition. Overall, we do not have a direct reference for Theorem 1, but we have no doubt that it is known.

Theorem 1. *Let f and G_i , $i = 1, \dots, m$, be convex. For each $\sigma > 0$, let x_σ be a solution of problem (2) with the barrier function defined in (3), and let $x_\sigma \rightarrow \bar{x}$ as $\sigma \rightarrow 0+$. Then the following assertions are valid:*

(i) *If the linear growth condition is satisfied at \bar{x} , then*

$$\|x_\sigma - \bar{x}\| = O(\sigma). \quad (18)$$

(ii) *If the quadratic growth condition is satisfied at \bar{x} , then*

$$\|x_\sigma - \bar{x}\| = O(\sigma^{1/2}). \quad (19)$$

The estimates obtained in Theorem 1 can be regarded as pure convergence rate estimates. Moreover, these estimates are sharp, even under LICQ, as demonstrated by the following simple examples.

Example 1. Let $n = m = 1$, $f(x) = x$, $G(x) = -x$ (linear functions). Clearly, $\bar{x} = 0$ is the unique solution of problem (1), and moreover, LICQ and FOSC (9) (hence, the linear growth condition) are satisfied at \bar{x} . It can be directly verified that for each $\sigma > 0$, the unique solution of subproblem (2) with the barrier function defined in (3) is given by $x_\sigma = \sigma$ and the estimate (18) is sharp. Moreover, $f(x_\sigma) = x_\sigma = \sigma$, and the estimate (17) is sharp as well.

Example 2. Let $n = m = 1$, $f(x) = x^2/2$, $G(x) = -x$ (convex functions). Note that the constraint in this example is the same as in Example 1. Evidently, $\bar{x} = 0$ is the unique solution of problem (1), and moreover, LICQ and SOSC (11) (hence, the quadratic growth condition, but not the linear growth condition!) are satisfied at \bar{x} . Note that the strict complementarity condition does not hold in this example. It can be directly verified that for each $\sigma > 0$, the unique solution of subproblem (2) with the barrier function defined in (3) is given by $x_\sigma = \sigma^{1/2}$ and the estimate (19) is sharp. Moreover, $f(x_\sigma) = x_\sigma^2/2 = \sigma/2$, and the estimate (17) is sharp as well.

In Proposition 1 and Theorem 1 we do not assume that MFCQ holds at \bar{x} , but MFCQ is implicitly subsumed in these results. As is well known, in the case of convex constraints, the condition $D^0 \neq \emptyset$ (called the *Slater CQ* [9]) is equivalent to MFCQ.

In order to proceed with the estimates for the nonconvex case, we need to assume explicitly that MFCQ holds at \bar{x} . In this case, it can be seen that the values μ_σ are bounded for all $\sigma > 0$ small enough. Indeed, suppose that there exists a sequence $\{\sigma_k\} \rightarrow 0+$ such that $\|\mu_{\sigma_k}\| \rightarrow \infty$. For each k , set $\bar{\mu}^k = \mu_{\sigma_k}/\|\mu_{\sigma_k}\|$. Then the sequence $\{\bar{\mu}^k\}$ has an accumulation point $\bar{\mu} \in \mathbb{R}^n \setminus \{0\}$. According to (12), for each $i \in \{1, \dots, m\} \setminus A$, it evidently holds that $(\mu^k)_i = \sigma/(-G_i(x_{\sigma_k})) \rightarrow 0$. Therefore, for such i , $(\bar{\mu}^k)_i = (\mu^k)_i/\|\mu_{\sigma_k}\| \rightarrow 0$ as $\sigma \rightarrow 0+$, where we have also taken into account that $\|\mu_{\sigma_k}\| \rightarrow \infty$, by the assumption. Hence, $\bar{\mu}_i = 0$ for all $i \in \{1, \dots, m\} \setminus A$. From (14), it follows that

$$\frac{1}{\|\mu_{\sigma_k}\|} f'(x_{\sigma_k}) + (G'(x_{\sigma_k}))^T \bar{\mu}^k = 0,$$

and by passing onto the limit along an appropriate subsequence, we obtain

$$(G'_A(\bar{x}))^T \bar{\mu}_A = 0, \quad \bar{\mu}_A \geq 0, \quad \bar{\mu}_A \neq 0,$$

which contradicts (the dual form of) MFCQ.

Another useful observation is that each accumulation point of μ_σ (as $\sigma \rightarrow 0+$) belongs to \mathcal{M} ; this follows from (12), (14).

We first consider the simpler case when the linear growth condition holds.

Theorem 2. *Assume that the linear growth condition and MFCQ hold at \bar{x} . For each $\sigma > 0$, let x_σ be a solution of problem (2) with the barrier function defined in (3), and let $x_\sigma \rightarrow \bar{x}$ as $\sigma \rightarrow 0+$. Then estimates (17) and (18) are valid.*

Proof. Using (13) and (14), and the above-mentioned fact that μ_σ is bounded as $\sigma \rightarrow 0+$, we obtain

$$\begin{aligned} f(x_\sigma) - f(\bar{x}) &= L(x_\sigma, \mu_\sigma) - f(\bar{x}) - \langle \mu_\sigma, G(x_\sigma) \rangle \\ &= L(x_\sigma, \mu_\sigma) - L(\bar{x}, \mu_\sigma) + \langle \mu_\sigma, G(\bar{x}) \rangle + m\sigma \\ &\leq - \left\langle \frac{\partial L}{\partial x}(x_\sigma, \mu_\sigma), x_\sigma - \bar{x} \right\rangle + m\sigma + o(\|x_\sigma - \bar{x}\|) \\ &= m\sigma + o(\|x_\sigma - \bar{x}\|). \end{aligned} \tag{20}$$

Estimate (18) follows from (8) and (20), while estimate (17) follows directly from (18) and (20). \square

The estimates obtained in Theorem 2 are sharp, even under LICQ, as is demonstrated by Example 1. Moreover, Theorem 2 actually extends assertion (i) of Theorem 1 to the nonconvex case.

The case when the weaker quadratic growth condition is assumed instead of the linear growth condition, is more complex. We consider this case next.

Theorem 3. *Assume that the quadratic growth condition and MFCQ hold at \bar{x} . For each $\sigma > 0$, let x_σ be a solution of problem (2) with the barrier function defined in (3), and let $x_\sigma \rightarrow \bar{x}$ as $\sigma \rightarrow 0+$. Then it holds that*

$$f(x_\sigma) \leq f(\bar{x}) + O\left(\sigma \left(\sum_{i \in A} |\ln(\mu_\sigma)_i| + 1\right)\right), \quad (21)$$

$$\|x_\sigma - \bar{x}\| = O\left(\sigma^{1/2} \left(\sum_{i \in A} |\ln(\mu_\sigma)_i| + 1\right)^{1/2}\right). \quad (22)$$

In particular,

$$f(x_\sigma) \leq f(\bar{x}) + O(\sigma |\ln \sigma|), \quad (23)$$

$$\|x_\sigma - \bar{x}\| = O(\sigma^{1/2} |\ln \sigma|^{1/2}). \quad (24)$$

Proof. For each $\sigma \geq 0$, consider the set

$$D_\sigma = \{x \in \mathbb{R}^n \mid G_i(x) \leq -\sigma, i = 1, \dots, m\}.$$

From MFCQ and Robinson's stability theorem [13], it follows that there exists $\tilde{x}_\sigma \in D_\sigma$ such that

$$\|\tilde{x}_\sigma - \bar{x}\| = O(\sigma). \quad (25)$$

Note that necessarily $\tilde{x}_\sigma \in D^0$ for each $\sigma > 0$. Using optimality of x_σ in problem (2), and (25), we obtain

$$\begin{aligned} f(x_\sigma) &= \varphi_\sigma(x_\sigma) + \sigma \sum_{i=1}^m \ln(-G_i(x_\sigma)) \\ &\leq \varphi_\sigma(\tilde{x}_\sigma) + \sigma \sum_{i=1}^m \ln(-G_i(x_\sigma)) \\ &= f(\tilde{x}_\sigma) + \sigma \sum_{i=1}^m (\ln(-G_i(x_\sigma)) - \ln(-G_i(\tilde{x}_\sigma))) \\ &\leq f(\bar{x}) + \sigma \sum_{i=1}^m \ln(G_i(x_\sigma)/G_i(\tilde{x}_\sigma)) + O(\sigma) \\ &\leq f(\bar{x}) + \sigma \sum_{i \in A} \ln(-G_i(x_\sigma)/\sigma) + O(\sigma) \\ &= f(\bar{x}) - \sigma \sum_{i \in A} \ln(\mu_\sigma)_i + O(\sigma), \end{aligned}$$

which implies estimate (21). Estimate (22) now follows from (10) and (21). The last two estimates in the assertion of the theorem are direct consequences of the first two, taking into account that for $i \in A$ and all $\sigma > 0$ sufficiently small, it holds that $\mu_\sigma = -\sigma/G_i(x_\sigma) \geq \sigma$, and hence, $|\ln(\mu_\sigma)_i| \leq |\ln \sigma|$. \square

Let us discuss the estimates obtained in Theorem 3. Clearly, estimates (21) and (22) are sharper than estimates (23) and (24), respectively. But the latter can be regarded as pure convergence rate estimates and are easier to use. Also, note that the right-hand side of (24) is $o(\sigma^\nu)$ for any $\nu \in (0, 1/2)$. Nevertheless, estimate (24) is somewhat weaker than (19), of course.

Estimate (22) cannot be regarded as a computable error bound of the form (5), because it contains the set A which depends on unknown \bar{x} . On the other hand, (22) implies (5) with

$$r(x_\sigma, \sigma) = \sigma^{1/2} \left(\sum_{i=1}^m |\ln(\mu_\sigma)_i| + 1 \right)^{1/2}, \quad (26)$$

which is computable. Of course, such bound is in general weaker than (22), because $(\mu_\sigma)_i \rightarrow 0$ as $\sigma \rightarrow 0+$ for each $i \in \{1, \dots, m\} \setminus A$; the latter follows from the above-mentioned fact that each accumulation point of μ_σ as $\sigma \rightarrow 0+$ belongs to \mathcal{M} . On the other hand, error bound (5) with $r(\cdot, \cdot)$ defined in (26) may be sharper than (24) (e.g., when $A = \{1, \dots, m\}$).

In the rest of this section, we are concerned with the possibilities to improve the estimates in Theorem 3 under some additional assumptions. Recall that this can be done for convex optimization problems: in this case, “ideal” estimates (17) and (19) hold, according to Proposition 1 and the assertion (ii) of Theorem 1.

Assume that f and G are twice differentiable, and their second derivatives are continuous at \bar{x} . Then, by direct computation, we obtain that for each $\sigma > 0$, it holds that

$$\frac{\partial^2 L}{\partial x^2}(x_\sigma, \mu_\sigma)[\xi, \xi] + \sigma \sum_{i=1}^m \frac{\langle G'_i(x_\sigma), \xi \rangle^2}{(G_i(x_\sigma))^2} = \varphi''_\sigma(x_\sigma)[\xi, \xi] \geq 0 \quad \forall \xi \in \mathbb{R}^n, \quad (27)$$

where the second-order necessary optimality conditions for problem (2) at x_σ are taken into account.

We start with an auxiliary estimate, which involves \bar{x} .

Proposition 2. *Assume that MFCQ holds at \bar{x} . For each $\sigma > 0$, let x_σ be a solution of problem (2) with the barrier function defined in (3), and let $x_\sigma \rightarrow \bar{x}$ as $\sigma \rightarrow 0+$. Then it holds that*

$$f(x_\sigma) \leq f(\bar{x}) + \frac{\sigma}{2} \sum_{i \in A} \frac{\langle G'_i(x_\sigma), x_\sigma - \bar{x} \rangle^2}{(G_i(x_\sigma))^2} + m\sigma + o(\|x_\sigma - \bar{x}\|^2). \quad (28)$$

Proof. Using (13), (14), and (27), and the above-mentioned fact that μ_σ is bounded as $\sigma \rightarrow 0+$, we obtain

$$\begin{aligned}
f(x_\sigma) - f(\bar{x}) &= L(x_\sigma, \mu_\sigma) - f(\bar{x}) - \langle \mu_\sigma, G(x_\sigma) \rangle \\
&= L(x_\sigma, \mu_\sigma) - L(\bar{x}, \mu_\sigma) + \langle \mu_\sigma, G(\bar{x}) \rangle + m\sigma \\
&\leq - \left\langle \frac{\partial L}{\partial x}(x_\sigma, \mu_\sigma), x_\sigma - \bar{x} \right\rangle - \frac{1}{2} \frac{\partial^2 L}{\partial x^2}(x_\sigma, \mu_\sigma)[x_\sigma - \bar{x}, x_\sigma - \bar{x}] \\
&\quad + m\sigma + o(\|x_\sigma - \bar{x}\|^2) \\
&= -\frac{1}{2} \varphi''_\sigma(x_\sigma)[x_\sigma - \bar{x}, x_\sigma - \bar{x}] + \frac{\sigma}{2} \sum_{i=1}^m \frac{\langle G'_i(x_\sigma), x_\sigma - \bar{x} \rangle^2}{(G_i(x_\sigma))^2} \\
&\quad + m\sigma + o(\|x_\sigma - \bar{x}\|^2) \\
&\leq \frac{\sigma}{2} \sum_{i \in A} \frac{\langle G'_i(x_\sigma), x_\sigma - \bar{x} \rangle^2}{(G_i(x_\sigma))^2} + m\sigma + o(\|x_\sigma - \bar{x}\|^2). \quad \square
\end{aligned}$$

Thus, the crucial question is the behavior of $\langle G'_i(x_\sigma), x_\sigma - \bar{x} \rangle / G_i(x_\sigma)$ for $i \in A$. If all these quantities are bounded as $\sigma \rightarrow 0+$, then (28) implies “ideal” estimate (17), and hence (19) under the quadratic growth condition. Clearly, this question is concerned with geometry of the feasible set and of the trajectory $\sigma \rightarrow x_\sigma$. In particular, it can be shown that if there exist $i \in A$ and a sequence $\{\sigma_k\} \rightarrow 0+$ such that $\langle G'_i(x_{\sigma_k}), x_{\sigma_k} - \bar{x} \rangle / G_i(x_{\sigma_k}) \rightarrow \infty$ as $k \rightarrow \infty$, then

$$G_i(x_{\sigma_k}) = o(\|x_{\sigma_k} - \bar{x}\|^2), \quad \langle G'_i(x_{\sigma_k}), x_{\sigma_k} - \bar{x} \rangle = O(\|x_{\sigma_k} - \bar{x}\|).$$

It turns out that the ideal estimates hold when the constraints are convex.

Theorem 4. *Assume that the quadratic growth condition and MFCQ hold at \bar{x} . For each $\sigma > 0$, let x_σ be a solution of problem (2) with the barrier function defined in (3), and let $x_\sigma \rightarrow \bar{x}$ as $\sigma \rightarrow 0+$. Let, in addition, G_i , $i \in A$, be convex. Then the estimates (17) and (19) are valid.*

Proof. According to the first-order convexity criterion for differentiable functions, for each $\sigma > 0$ it holds that

$$-G_i(x_\sigma) = G_i(\bar{x}) - G_i(x_\sigma) \geq -\langle G'_i(x_\sigma), x_\sigma - \bar{x} \rangle \quad \forall i \in A.$$

Thus for $i \in A$, quantities $\langle G'_i(x_\sigma), x_\sigma - \bar{x} \rangle / G_i(x_\sigma)$ are bounded (above by one), and the needed estimates follow from Proposition 2. \square

Theorem 4 extends assertion (ii) of Theorem 1 to the case where the objective function need not be convex.

3 Error Estimates for the Inverse-Barrier Method

Throughout this section, φ_σ is defined with the inverse barrier (4). The development is similar to that of Section 2, except that we obtain (computable)

estimates in terms of σ and x_σ , rather than convergence rates depending on σ only.

For each $\sigma > 0$, denote

$$\mu_\sigma = \sigma(1/(G_1(x_\sigma))^2, \dots, 1/(G_m(x_\sigma))^2) > 0. \quad (29)$$

By direct computation,

$$\langle \mu_\sigma, G(x_\sigma) \rangle = \sigma \sum_{i=1}^m 1/G_i(x_\sigma), \quad (30)$$

and by the first-order necessary optimality conditions for problem (2) at x_σ , we also have condition (14) (but with μ_σ defined in (29)).

If (1) is a convex minimization problem, we have the following.

Proposition 3. *Let f and G_i , $i = 1, \dots, m$, be convex. For $\sigma > 0$, let x_σ be a solution of problem (2) with the barrier function defined in (4). Then it holds that*

$$f(x_\sigma) \leq \inf_{x \in D} f(x) - \sigma \sum_{i=1}^m 1/G_i(x_\sigma).$$

Proof. The assertion follows from observing that (x_σ, μ_σ) is a feasible point for the Wolfe dual of (1), and using the weak duality relation. \square

In our setting, it therefore holds that

$$f(x_\sigma) \leq f(\bar{x}) + O\left(\sigma \sum_{i=1}^m (-1/G_i(x_\sigma))\right). \quad (31)$$

Recall that in the convex case, MFCQ is equivalent to our standing assumption that $D^0 \neq \emptyset$. By the same argument as in Section 2 (but using (29) instead of (12)), it can be shown that the values μ_σ are bounded for all $\sigma > 0$ small enough, and that each accumulation point of μ_σ (as $\sigma \rightarrow 0+$) belongs to \mathcal{M} . Then (30) implies that

$$\sigma \sum_{i=1}^m (-1/G_i(x_\sigma)) \rightarrow 0 \text{ as } \sigma \rightarrow 0+,$$

which shows that the estimate (31) is meaningful.

From Proposition 3, we immediately obtain the following result.

Theorem 5. *Let f and G_i , $i = 1, \dots, m$, be convex. For each $\sigma > 0$, let x_σ be a solution of problem (2) with the barrier function defined in (4), and let $x_\sigma \rightarrow \bar{x}$ as $\sigma \rightarrow 0+$. Then the following assertions are valid:*

(i) If the linear growth condition is satisfied at \bar{x} , then

$$\|x_\sigma - \bar{x}\| = O\left(\sigma \sum_{i=1}^m (-1/G_i(x_\sigma))\right). \quad (32)$$

(ii) If the quadratic growth condition is satisfied at \bar{x} , then

$$\|x_\sigma - \bar{x}\| = O\left(\sigma^{1/2} \sum_{i=1}^m (-1/G_i(x_\sigma))^{1/2}\right). \quad (33)$$

Examples 1 and 2 can be used in order to demonstrate that the estimates obtained in Theorem 5 are sharp; we omit the details.

We proceed with the estimates for the nonconvex case, explicitly assuming that MFCQ holds at \bar{x} .

Theorem 6. *Assume that the linear growth condition and MFCQ hold at \bar{x} . For each $\sigma > 0$, let x_σ be a solution of problem (2) with the barrier function defined in (4), and let $x_\sigma \rightarrow \bar{x}$ as $\sigma \rightarrow 0+$. Then estimates (31) and (32) are valid.*

Proof. The proof is similar to that of Theorem 2, but using (30) instead of (13). \square

We note that a counterpart of Theorem 3 does not hold for the inverse barrier.

Assume now that f and G are twice differentiable, and their second derivatives are continuous at \bar{x} . By direct computation, we obtain that for each $\sigma > 0$, it holds that

$$\frac{\partial^2 L}{\partial x^2}(x_\sigma, \mu_\sigma)[\xi, \xi] - 2\sigma \sum_{i=1}^m \frac{\langle G'_i(x_\sigma), \xi \rangle^2}{(G_i(x_\sigma))^3} = \varphi''_\sigma(x_\sigma)[\xi, \xi] \geq 0 \quad \forall \xi \in \mathbb{R}^n, \quad (34)$$

where the second-order necessary optimality conditions for problem (2) at x_σ are taken into account.

Following the lines of the proofs of Proposition 2 and Theorem 4, we obtain the next two results.

Proposition 4. *Assume that MFCQ holds at \bar{x} . For each $\sigma > 0$, let x_σ be a solution of problem (2) with the barrier function defined in (4), and let $x_\sigma \rightarrow \bar{x}$ as $\sigma \rightarrow 0+$. Then it holds that*

$$f(x_\sigma) \leq f(\bar{x}) + \sigma \sum_{i \in A} \frac{1}{-G_i(x_\sigma)} \left(\frac{\langle G'_i(x_\sigma), x_\sigma - \bar{x} \rangle^2}{(G_i(x_\sigma))^2} + 1 \right) + o(\|x_\sigma - \bar{x}\|^2).$$

Theorem 7. *Assume that the quadratic growth condition and MFCQ hold at \bar{x} . For each $\sigma > 0$, let x_σ be a solution of problem (2) with the barrier function defined in (4), and let $x_\sigma \rightarrow \bar{x}$ as $\sigma \rightarrow 0+$. Let, in addition, G_i , $i \in A$, be convex. Then the estimates (31) and (33) are valid.*

4 Concluding Remarks

We presented computable error bound estimates and convergence rate results for some interior penalty methods. Our assumptions are essentially the Mangasarian-Fromovitz constraint qualification and the linear or quadratic growth condition (in this setting, the latter are equivalent to the first-order or second-order sufficient optimality conditions, respectively).

Some of the estimates are shown to be sharp. But at this time, it is an open question whether the estimates for the log-barrier method given under the Mangasarian-Fromovitz constraint qualification and the quadratic growth condition (Theorem 3) are sharp.

References

1. L. Bittner. Eine Verallgemeinerung des Verfahrens des logarithmischen Potentials von Frisch für nichtlineare Optimierungsprobleme. In: A. Pekora (ed.), *Colloquium on Applications of Mathematics to Economics*, Budapest, 1963, Akademiai Kiado. Publishing House of the Hungarian Acad. of Sciences, 1965.
2. J.F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer-Verlag, New York, 2000.
3. A.V. Fiacco and G.P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. John Wiley & Sons, New York, 1968.
4. A. Fischer. Local behaviour of an iterative framework for generalized equations with nonisolated solutions. *Math. Program.* 94: 91–124, 2002.
5. C. Grossman, D. Klatté, and B. Kummer. Convergence of primal-dual solutions for the nonconvex log-barrier method without LICQ. *Kybernetika* 40: 571–584, 2004.
6. W.W. Hager and M.S. Gowda. Stability in the presence of degeneracy and error estimation. *Math. Program.* 85: 181–192, 1999.
7. A.F. Izmailov and M.V. Solodov. Computable primal error bounds based on the augmented Lagrangian and Lagrangian relaxation algorithms. Preprint A 2004/303, IMPA, Rio de Janeiro, 2004.
8. A.F. Izmailov and M.V. Solodov. Karush-Kuhn-Tucker systems: regularity conditions, error bounds and a class of Newton-type methods. *Math. Program.* 95: 631–650, 2003.
9. O.L. Mangasarian. *Nonlinear Programming* McGraw-Hill, New York, 1969
10. R. Mifflin. On the convergence of the logarithmic barrier function method. In: F.A. Lootsma (ed.), *Numerical Methods for Nonlinear Optimization*. Academic Press, London, 1972.
11. R. Mifflin. Convergence bounds for nonlinear programming algorithms. *Math. Program.* 8: 251–271, 1975.
12. J.S. Pang. Error bounds in mathematical programming. *Math. Program.* 79: 299–332, 1997.
13. S.M. Robinson. Stability theorems for systems of inequalities, Part II: differentiable nonlinear systems. *SIAM J. Numer. Anal.* 13: 497–513, 1976.
14. S.J. Wright and D. Orban. Properties of the log-barrier function on degenerate nonlinear programs. *Math. Oper. Res.* 27: 585–613, 2002.