J. Austral. Math. Soc. **21** (Series B) (1980), 305–329

A MODIFIED BARRIER FUNCTION METHOD WITH IMPROVED RATE OF CONVERGENCE FOR DEGENERATE PROBLEMS

KRISORN JITTORNTRUM and M. R. OSBORNE

(Received 9 November 1978)

(Revised 19 April 1979)

Abstract

In a previous paper the authors have shown that the classical barrier function has an $O(r)$ rate of convergence unless the problem is degenerate when it reduces $O(r^{\frac{1}{2}})$. In this paper a modified barrier function algorithm is suggested which does not suffer from this problem. It turns out to have superior scaling properties which make it preferable to the classical algorithm, even in the nondegenerate case, if extrapolation is to be used to accelerate convergence.

1. Introduction

In a recent paper [4] we have considered the solution of the mathematical programming problem (MPP),

$$\min_{\mathbf{x}\in S} f(\mathbf{x}): \quad S = \{\mathbf{x};\, g_i(\mathbf{x}) \geqslant 0,\ i = 1, 2, \ldots, m\}, \tag{1.1}$$

where f and the g_i, $i = 1, 2, \ldots, m$, are appropriately smooth functions on $R^n \to R$, by means of the sequential minimization of the barrier function (classical barrier function),

$$B(\mathbf{x}, r) = f(\mathbf{x}) - r \sum_{i=1}^{m} \log(g_i(\mathbf{x})), \tag{1.2}$$

for r taking values $r_1 > r_2 > \ldots > r_k > \ldots$ and $\lim_{k\to\infty} r_k = 0$. Let $\mathbf{x}(r_k)$ be the exact minimum of (1.2) produced by some algorithmic procedure for $r = r_k$. We assume that the minimum exists and is well defined. Then it is well known that the limit points of the sequence $\{\mathbf{x}(r_k)\}$ are solutions of the MPP under very general conditions [1]. In particular, these limit points are solutions of the MPP if the following propositions hold [1].

PROPOSITION 1. (*First-order necessary conditions.*) *Provided an appropriate regularity condition is satisfied by the feasible region S then a necessary condition for* $\mathbf{x}^* \in S$ *to be a solution of the MPP is that there exist multipliers* u_i^*, $i = 1, 2, \ldots, m$, *satisfying the Kuhn–Tucker conditions*:

$$\text{(a)}\quad \nabla f(\mathbf{x}^*) - \sum_{i=1}^{m} u_i^* \nabla g_i(\mathbf{x}^*) = 0, \tag{1.3a}$$

$$\text{(b)}\quad u_i^* \geqslant 0, \quad u_i^* g_i(\mathbf{x}^*) = 0, \quad i = 1, 2, \ldots, m. \tag{1.3b}$$

Here we specialize the regularity condition to be the condition that if $\mathbf{x} \in S$ *then the set of vectors* $\nabla g_i(\mathbf{x})$ *corresponding to the constraints satisfying* $g_i(\mathbf{x}) = 0$ *be linearly independent. Strictly this condition is required only when* $\mathbf{x} = \mathbf{x}^*$.

To specify the second proposition we define the Lagrangian function for the MPP,

$$L(\mathbf{x}, \mathbf{u}) = f(\mathbf{x}) - \sum_{i=1}^{m} u_i g_i(\mathbf{x}). \tag{1.4}$$

PROPOSITION 2. (*Second-order sufficiency conditions.*) *Let T denote the set*

$$T = \{\mathbf{t};\ \nabla g_i(\mathbf{x}^*)\mathbf{t} = 0, \quad \textit{for all } i \textit{ such that } u_i^* > 0\}.$$

If the Kuhn–Tucker conditions are satisfied at $\mathbf{x}^*$, *and if there exists* $m > 0$ *such that (for some appropriate vector norm)*

$$\mathbf{t}^{\mathrm{T}} \nabla_x^2 L(\mathbf{x}^*, \mathbf{u}^*)\mathbf{t} \geqslant m \|\mathbf{t}\|^2, \quad \textit{for all } \mathbf{t} \in T, \tag{1.5}$$

then there exists an open neighbourhood N of $\mathbf{x}^*$ *in S such that, if* $\mathbf{x} \in N$ *and* $\mathbf{x} \neq \mathbf{x}^*$, *then* $f(\mathbf{x}) > f(\mathbf{x}^*)$.

Now let $\mathbf{x}^*$ be a limit point of the sequence of barrier function minimizations $\{\mathbf{x}(r_k)\}$. We define $\mathbf{x}^*$ to be a *regular local solution* if the first-order necessary and second-order sufficiency conditions are satisfied at $\mathbf{x}^*$.

REMARK 1.1. If $\mathbf{x}^*$ is a regular local solution of the MPP then
(a) the Kuhn–Tucker conditions hold at $\mathbf{x}^*$ and the multipliers u_i^* are unique,
(b) the sequence of values $\{r_k/g_i(\mathbf{x}(r_k))\} \to u_i^*$ for each i, and
(c) $\mathbf{x}^*$ is an isolated minimum of the MPP.

The main results of [3] give convergence rates for $\{\mathbf{x}(r_k)\}$ and $\{\mathbf{u}(r_k)\}$, where $u_i(r_k) = r_k/g_i(\mathbf{x}(r_k))$, as $r \to 0$. It is shown that

$$\|\mathbf{x}(r_k) - \mathbf{x}^*\| = O(r_k) \quad \text{and} \quad \|\mathbf{u}(r_k) - \mathbf{u}^*\| = O(r_k)$$

if and only if the MPP is nondegenerate at $\mathbf{x}^*$. If the MPP is degenerate at $\mathbf{x}^*$, then the rate of convergence drops to exactly $O(r_k^{\frac{1}{2}})$. To explain the nondegeneracy condition let

$$I = \{i;\ g_i(\mathbf{x}^*) = 0\}.$$

If $i \in I$ then the corresponding constraint $g_i(\mathbf{x})$ is said to be active at $\mathbf{x}^*$. The complement of I with respect to $\{1, 2, \ldots, m\}$ is

$$R = \{1, 2, \ldots, m\} - I.$$

If $i \in R$ then $g_i(\mathbf{x}^*) > 0$. We also write $I = I_1 \cup I_2$ where

$$I_1 = \{i;\ g_i(\mathbf{x}^*) = 0,\ u_i^* > 0\}$$

and

$$I_2 = \{i;\ g_i(\mathbf{x}^*) = 0,\ u_i^* = 0\}.$$

DEFINITION. *The MPP is* degenerate *at* $\mathbf{x}^*$ *if* $I_2 \neq \varnothing$.

Thus the results cited above show that degeneracy causes a serious degradation in the performance of the barrier function algorithm, and this shows up geometrically by the solution trajectory defined by $\{\mathbf{x}(r_k)\}$ being forced to approach $\mathbf{x}^*$ ultimately tangentially to the constraint surfaces $g_i(\mathbf{x}) = 0$, $i \in I_1$.

Note that a degenerate constraint is redundant in an important sense as the first-order necessary conditions are unchanged if this constraint is just ignored, and as degeneracy does not affect the second-order sufficiency conditions. However, it is an indication that the property of membership of the active constraint set is extremely sensitive to perturbations of the problem data. A characteristic of the classical barrier function method is that there is no discrimination between members of the constraint set. This results in several undesirable properties: (i) it limits the usefulness of extrapolation procedures for improving the numerical performance of the algorithm on near-degenerate problems; (ii) the rate of convergence is reduced to $O(r^{\frac{1}{2}})$ for degenerate problems; and (iii) the inability to identify inactive constraints clearly slows down convergence in situations where the solution $\mathbf{x}^*$ is properly in the interior of the feasible region.

The main aim of this paper is to provide a sound theoretical basis for a method which modifies the barrier objective function in such a way as to overcome these problems. This method was suggested in the report [6], and the key feature is the use of the best current estimate of the Lagrange multipliers to weight the constraints in the current minimization. The modified barrier function thus has the form

$$M(\mathbf{x}, r_k \mathbf{u}^{(k-1)}) = f(\mathbf{x}) - r_k \sum_{i \in I} u_i^{(k-1)} \log(g_i(\mathbf{x})), \tag{1.6}$$

where $\mathbf{u}^{(0)}$, r_1 and $c > 1$ are given, $r_k = r_{k-1}/c$ for $k = 1, 2, \ldots$, and where

$$u_i^{(k)} = r_k u_i^{(k-1)}/g_i(\mathbf{x}^{(k)}), \quad i \in I, \tag{1.7}$$

gives the multiplier estimates after the kth minimization. Note that the modified barrier function algorithm is strictly a sequential algorithm as the minimum $\mathbf{x}^{(k)}$ depends not only on r_k but also on the previous minimization through $\mathbf{u}^{(k-1)}$. A consequence is that the trajectory analysis given in [3] no longer applies. However, this problem can be overcome in the nondegenerate case by considering the modified barrier function $M(\mathbf{x}, \mathbf{s})$ with $\mathbf{s} = r\phi(r)$, where $\phi(r)$ denotes a differentiable vector-valued function of r. This form of the modified barrier function is of interest because a smooth trajectory can be defined for the successive minima, and because the special function ϕ, given by

$$\phi(r) = \mathbf{u}(cr), \tag{1.8}$$

can be constructed in the nondegenerate case. However, the modified barrier function is perhaps of most interest in the degenerate case. Consider the following example.

EXAMPLE 1.1,

$$\begin{aligned} &\text{Minimize} && f(\mathbf{x}) = x_2, \\ &\text{subject to} && g_1(\mathbf{x}) = x_2 - x_1^2 \geqslant 0, \\ &\text{and} && g_2(\mathbf{x}) = x_1 \geqslant 0. \end{aligned}$$

The solution is $\mathbf{x}^* = (0,0)$, $\mathbf{u}^* = (1,0)$, and the problem is degenerate as both constraints are active. The modified barrier function has the form

$$M(\mathbf{x}, r\phi(r)) = x_2 - r\phi_1(r)\log(x_2 - x_1^2) - r\phi_2(r)\log(x_1), \tag{1.9}$$

and, at a minimum $\mathbf{x} = \mathbf{x}(r)$ of $M(\mathbf{x}, r\phi(r))$,

$$(0,0) = (0,1) - \frac{r\phi_1(r)}{x_2 - x_1^2}(-2x_1, 1) - \frac{r\phi_2(r)}{x_1}(1,0). \tag{1.10}$$

We consider two cases:

(i) $\phi_i(r) = 1$, $i = 1, 2$. This corresponds to the classical barrier function. We have

$$x_1(r) = (r/2)^{\frac{1}{2}}, \quad x_2(r) = 3r/2, \tag{1.11}$$

$$u_1(r) = 1, \quad u_2(r) = (2r)^{\frac{1}{2}}, \tag{1.12}$$

showing the predicted $O(r^{\frac{1}{2}})$ convergence.

(ii) $\phi_1(r) = 1$, $\phi_2(r) = 2cr$ for $c \geqslant 1$. We have

$$x_1(r) = cr, \quad x_2(r) = r + c^2r^2, \tag{1.13}$$

$$u_1(r) = 1, \quad u_2(r) = 2cr, \tag{1.14}$$

Thus the convergence rate for this particular choice of $\phi(r)$ is $O(r)$. In this case ϕ

https://doi.org/10.1017/S033427000000240X Published online by Cambridge University Press

has been chosen such that

$$\phi(r) = (1, 2cr) = \mathbf{u}(cr), \tag{1.15}$$

Now consider use of the modified barrier function algorithm (1.6), (1.7) to solve the problem. At each iteration we minimize

$$M(\mathbf{x}, r_k \mathbf{u}^{(k-1)}) = x_2 - r_k u_1^{(k-1)} \log(x_2 - x_1^2) - r_k u_2^{(k-1)} \log(x_1), \tag{1.16}$$

giving

$$u_1^{(k)} = 1, \quad u_2^{(k)} = 2cr_k \left[\frac{u_2^{(0)}}{2r_1 c^2} \right]^{(\frac{1}{2})^k}. \tag{1.17}$$

By comparing $\mathbf{u}^{(k)}$ given by (1.14) and (1.17) it follows that if the modified algorithm is started with $u_2^{(0)} = 2r_1 c^2$ then the resulting multipliers are the same. The modified algorithm is in effect producing the continuous solution (1.13) for various values of r_k and second-order extrapolation as described in [4] gives the exact solution $\mathbf{x}(0) = (0, 0)$. However, usually it cannot be expected that the initial multiplier will be correct. In this example, the ratio of the values of $u_2^{(k)}$ from equations (1.17) and (1.14) is $[u_2^{(0)}/2r_1 c^2]^{(\frac{1}{2})^k}$ and tends to 1 as k becomes large, showing that the modified algorithm is convergent essentially independent of the value of $u_2^{(0)}$. However, this ratio can be interpolated as a smooth function of r only if $\log 2/\log c$ is integral. Thus, while $O(r)$ extrapolation is possible, the largest value of c for which higher order extrapolation appears feasible is $c = 2$.

The behaviour in this example seems typical. In this connection our major result is that the modified barrier function algorithm gives $O(r)$ convergence for degenerate problems and thus supports first-order extrapolation. Higher order extrapolation is possible for nondegenerate problems but has not been established in the degenerate case. Both the above example and the numerical results presented in Section 4 appear to indicate that it is unlikely to be worthwhile in the degenerate case.

Thus we are able to provide both a theoretical basis and supporting numerical evidence for the use of the modified barrier function algorithm. Our main conclusion is that it is clearly superior to the classical algorithm and so is definitely attractive for MPP's *in which the problem functions are not defined outside the feasible region.*† This superiority shows up even in nondegenerate problems. This is because the performance of extrapolation procedures applied to values obtained by minimizing the classical barrier function is known to depend strongly on the range of values of the Lagrange multipliers [4]. We show in the next section that

† The property of maintaining feasibility is the major advantage of barrier function methods. If this is not important then there are better methods, at least in the non-degenerate case (see, for example, Jittorntrum [3] and Powell [7]). If the problem is degenerate then the question is more complicated. A discussion is given in [2].

these difficulties do not occur in the modified algorithm since the weighting of the barrier terms using the latest estimates of the Lagrange multipliers is in a sense asymptotically optimal providing the scaling of the constraints is not to disparate. (For the result to hold strictly we should have $\| \nabla g_i(\mathbf{x}^*) \| = 1$, $i \in I$.)

2. Properties of the modified barrier function algorithm

In this section we derive properties of the modified barrier function algorithm defined by equations (1.6) and (1.7). It turns out that there are important differences between the degenerate and nondegenerate cases and we consider these separately. We assume that $\mathbf{x}^*$ is a regular local solution of the MPP and that the sequences $\{\mathbf{x}^{(k)}\}$ and $\{\mathbf{u}^{(k)}\}$ generated by the modified algorithm converge to $\mathbf{x}^*$ and $\mathbf{u}^*$ respectively.

DEFINITION. *Let $A(r)$ be a continuous function of r, $\{r_k\}$ a strictly decreasing sequence tending to zero, and $\{A^{(k)}\}$ a sequence of values. Then the sequence $\{A^{(k)}\}$ is said to converge* super fast *to $A(r)$ if there exist bounded constants μ, ρ such that*

$$\| A^{(k)} - A(r_k) \| \leqslant \mu \rho^k \prod_{i=1}^{k} r_i, \quad k = 1, 2, \ldots. \tag{2.1}$$

With the usual barrier function, $g_i \to g_i(\mathbf{x}^*) > 0$ for $i \in R$, while the corresponding multiplier is given by $u_i = r/g_i \to 0$ with r. Thus eventually the contribution from the inactive constraints fades out of the barrier function as r gets small. In contrast, the contribution of the inactive constraints gets small super fast in the modified algorithm. This is the content of the following theorem.

THEOREM 2.1. *Assume that $\mathbf{x}^*$ is a regular local solution of the MPP, and that $g_i(\mathbf{x}^*) > 0$; then the sequence of multiplier estimates $\{u_i^{(k)}\}$ generated by the modified algorithm converges super fast to zero.*

PROOF. We have, by (1.7),

$$\begin{aligned} u_i^{(k)} &= r_k u_i^{(k-1)} / g_i(\mathbf{x}^{(k)}) \\ &= r_k r_{k-1} u_i^{(k-2)} / [g_i(\mathbf{x}^{(k)}) g_i(\mathbf{x}^{(k-1)})] \end{aligned}$$

and, proceeding recursively, it follows that

$$u_i^{(k)} \leqslant u_i^{(0)} \rho_i^k \prod_{j=1}^{k} r_j, \tag{2.2}$$

where ρ_i is a bound for $1/g_i(\mathbf{x}^{(k)})$, $k = 1, 2, \ldots$.

REMARK 2.1. This result permits the inactive constraints to be readily identified, as their contribution to the objective function becomes negligible after a few iterations. In particular, if all constraints are inactive at the minimum, the successive minima of the modified algorithm converge super fast to this minimum. This should be compared with the $O(r)$ rate of convergence of the classical algorithm in this case.

In the nondegenerate case we can extend the trajectory analysis given in [4] to provide information on the modified algorithm. We recall that the points minimizing (1.2) as $r \to 0$ lie on the smooth trajectory given by the solution of the system of differential equations:

$$J(\mathbf{x},\mathbf{u})\begin{bmatrix} \dfrac{d\mathbf{x}}{dr} \\ \dfrac{d\mathbf{u}}{dr} \end{bmatrix} = \begin{bmatrix} \nabla_x^2 L & -\nabla g_1^{\mathrm{T}} & \cdots & -\nabla g_m^{\mathrm{T}} \\ u_1 \nabla g_1 & g_1 & & \\ \vdots & & \ddots & \\ u_m \nabla g_m & & & g_m \end{bmatrix} \begin{bmatrix} \dfrac{d\mathbf{x}}{dr} \\ \dfrac{d\mathbf{u}}{dr} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \end{bmatrix}. \tag{2.3}$$

When the problem is nondegenerate, $J(\mathbf{x}^*, \mathbf{u}^*)$ is nonsingular so that repeated differentiation of (2.3) determines derivatives of $\mathbf{x}$ and $\mathbf{u}$ to any desired order. It follows, in particular, that $\| \mathbf{x}(r) - \mathbf{x}^* \| = O(r)$ and $\| \mathbf{u}(r) - \mathbf{u}^* \| = O(r)$. The trajectory analysis for the modified barrier function is complicated because it is necessary to guess the $u_i^{(0)}$ at the first step to start the process. However, we can show in the nondegenerate case that there is a well-defined smooth trajectory, and that the results of our sequence of minimizations tend to this trajectory super fast.

THEOREM 2.2. (1) *Let* $\mathbf{x}^*$ *be a regular local solution of a nondegenerate MPP, and assume that all the problem functions are* l *times continuously differentiable. Then, for given* $c > 1$, *there exists an* $l-1$ *times continuously differentiable vector function* $\hat{\phi}(r)$, *uniquely determined up to the first* l *terms of its Taylor expansion about* $r = 0$, *such that*

$$\hat{\phi}(r) = \hat{\mathbf{u}}(cr) + O(r^l), \tag{2.4}$$

where

$$\hat{u}_i(r) = r\hat{\phi}_i(r)/g_i(\hat{\mathbf{x}}(r)), \quad i = 1, 2, \ldots, m, \tag{2.5}$$

are the multiplier estimates, and $\hat{\mathbf{x}}(r)$ *is the minimizer of* $M(\mathbf{x}, r\hat{\phi}(r))$.

(2) *We conclude that* $\hat{\phi}(r) = \hat{\mathbf{u}}(cr)$ *exists for problems that are sufficiently smooth, and that the resulting modified barrier function defines a smooth solution trajectory* $(\hat{\mathbf{x}}(r), \hat{\mathbf{u}}(r))$. *Let* $r_k = r_{k-1}/c$, $k = 2, 3, \ldots$, *and* $\{\mathbf{x}^{(k)}\}$ *be the corresponding sequence of points generated by the modified barrier function algorithm. Then* $\{\mathbf{x}^{(k)}\}$ *tends super fast to* $\hat{\mathbf{x}}(r)$.

PROOF. Let $\hat{\phi}(r)$ have the form

$$\hat{\phi}(r) = \mathbf{u}^* + c\mathbf{a}_1 r + \ldots + c^{l-1}\mathbf{a}_{l-1}\frac{r^{l-1}}{(l-1)!} + O(r^l). \tag{2.7}$$

Then it follows from (2.4) that $\hat{\mathbf{u}}(r)$ is given by

$$\hat{\mathbf{u}}(r) = \mathbf{u}^* + \mathbf{a}_1 r + \ldots + \mathbf{a}_{l-1}\frac{r^{l-1}}{(l-1)!} + O(r^l). \tag{2.8}$$

We show that the coefficients $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{l-1}$ are well determined by the problem.

At the minimum of $M(\mathbf{x}, r\hat{\phi}(r))$ we have

$$\nabla f(\hat{\mathbf{x}}(r)) - \sum_{i=1}^{m} \hat{u}_i(r)\nabla g_i(\hat{\mathbf{x}}(r)) = 0, \tag{2.9a}$$

and

$$\hat{u}_i(r) g_i(\hat{\mathbf{x}}(r)) = r\hat{\phi}_i(r), \quad i = 1, 2, \ldots, m. \tag{2.9b}$$

Differentiating the equations (2.9) gives the system of differential equations

$$J(\hat{\mathbf{x}}(r), \hat{\mathbf{u}}(r))\begin{bmatrix} \dfrac{d\hat{\mathbf{x}}(r)}{dr} \\ \dfrac{d\hat{\mathbf{u}}(r)}{dr} \end{bmatrix} = \begin{bmatrix} 0 \\ \dfrac{d}{dr}(r\hat{\phi}(r)) \end{bmatrix}. \tag{2.10}$$

Because the problem is nondegenerate, $J(\mathbf{x}^*, \mathbf{u}^*)$ is nonsingular. Thus, setting $r = 0$ in (2.10),

$$\begin{bmatrix} \dfrac{d\hat{\mathbf{x}}}{dr}(0) \\ \mathbf{a}_1 \end{bmatrix} = J(\mathbf{x}^*, \mathbf{u}^*)^{-1}\begin{bmatrix} 0 \\ \mathbf{u}^* \end{bmatrix}. \tag{2.11}$$

Differentiating (2.10) with respect to r gives

$$\begin{bmatrix} \dfrac{d^2\hat{\mathbf{x}}}{dr^2} \\ \dfrac{d^2\hat{\mathbf{u}}}{dr^2} \end{bmatrix} = \begin{bmatrix} 0 \\ \dfrac{d^2}{dr^2}(r\hat{\phi}) \end{bmatrix} - \frac{dJ}{dr}\begin{bmatrix} \dfrac{d\hat{\mathbf{x}}}{dr} \\ \dfrac{d\hat{\mathbf{u}}}{dr} \end{bmatrix}, \tag{2.12}$$

and setting $r = 0$ in (2.12) we find that $\mathbf{a}_2$ is given in terms of known quantities. Proceeding in this fashion, we find successively $\mathbf{a}_3, \ldots \mathbf{a}_{l-1}$.

To prove part 2 it is only necessary to show that the sequence of multiplier estimates $\{\mathbf{u}^{(k)}\}$ given by the modified algorithm (1.6), (1.7) converges super fast to $\hat{\mathbf{u}}(r)$. (This follows, for example, from (2.10) which shows that perturbations in $\hat{\phi}$ lead to perturbations of the same order $d\mathbf{x}/dr$ provided r is small enough and the problem is nondegenerate.) Now consider the family of modified barrier functions $M(\mathbf{x}, \mathbf{s}^{(k)}(r, \theta))$ where

$$s_i^{(k)}(r, \theta_0) = r u_i^{(k-1)},$$
$$s_i^{(k)}(r, \theta_1) = r \hat{u}\,(r_{k-1}),$$

and $\mathbf{s}^{(k)}$ is a smooth function of the parameter θ. For $\|\mathbf{s}^{(k)}\|$ small enough, $\mathbf{u}(\mathbf{s}^{(k)}(r, \theta))$ can be defined by the leading terms of its multivariate Taylor series which can be evaluated readily at $r = 0$. Provided $r = r_k$ is small enough, an application of the mean value theorem gives

$$\begin{aligned} u_i^{(k)} - \hat{u}_i(r_k) &= u_i(\mathbf{s}^{(k)}(r_k, \theta_0)) - u_i(\mathbf{s}^{(k)}(r_k, \theta_1)) \\ &= \sum_{j=1}^{m} \overline{\frac{\partial u_i}{\partial s_j^{(k)}}} \{s_j^{(k)}(r_k, \theta_0) - s_j^{(k)}(r_k, \theta_1)\} \\ &= r_k \sum_{j=1}^{m} \overline{\frac{\partial u_i}{\partial s_j^{(k)}}} \{u_j^{(k-1)} - \hat{u}_j(r_{k-1})\}, \end{aligned} \tag{2.13}$$

where the bar denotes that appropriate mean values are taken. It follows from (2.13) that $\{\hat{\mathbf{u}}^{(k)}\}$ converges super fast to $\hat{\mathbf{u}}(r)$.

REMARK 2.2. A consequence of the super fast convergence is that extrapolation can be used to improve the estimates of $\mathbf{x}^*$ obtained from the modified barrier function algorithm. The first few points calculated will be most affected by the error in the initial guess $\mathbf{u}^{(0)}$, and in general the first at least should be omitted from the extrapolation process.

REMARK 2.3. For the classical barrier function algorithm the convergence of $\mathbf{x}^{(k)}$ to $\mathbf{x}^*$ is such that asymptotically

$$g_j(\mathbf{x}^{(k)}) \sim \frac{r_k}{u_j^*}, \quad j \in I. \tag{2.14}$$

Now the quantity $\max_{i \in I} |g_i(\mathbf{x}^{(k)})|$ gives an indication of the convergence of $\mathbf{x}^{(k)}$ to $\mathbf{x}^*$, and the quantity $\min_{i \in I} |g_i(\mathbf{x}^{(k)}|$ indicates the degree of difficulty of the unconstrained minimization because the corresponding barrier function term gets large if any $g_i(\mathbf{x})$ gets small. It follows from (2.14) that the ratio of these two quantities is proportional to γ, the measure of degeneracy introduced in [4]. Thus the size of γ in the classical barrier function algorithm corresponds to the property that the approach to $\mathbf{x}^*$ is such that the distance of $\mathbf{x}^{(k)}$ from each constraint $g_i(\mathbf{x}) = 0$, $i \in I$, is inversely proportional to the size of the multiplier u_i^*. However, for the modified algorithm,

$$g_i(\mathbf{x}^{(k)}) = \frac{r_k u_i^{(k-1)}}{u_i^{(k)}} \sim r_k, \quad i \in I. \tag{2.15}$$

Now the distance from the ith constraint is asymptotic to $r_k/\|\nabla g_i\|$ so that, if the $\|\nabla g_i\|$ are not too disparate, the effect of the size of γ is removed.

In the degenerate case it follows from (2.3) that $J(\mathbf{x}^*, \mathbf{u}^*)$ has zero rows corresponding to the degenerate constraints and so is singular. For the classical barrier function this results in $\|\mathbf{x}(r)-\mathbf{x}^*\| = O(r^{\frac{1}{2}})$ and $\|\mathbf{u}(r)-\mathbf{u}^*\| = O(r^{\frac{1}{2}})$. Thus if we want to follow the development of the modified barrier function given for the nondegenerate case we must first determine what conditions must be satisfied by $\phi(r)$ to ensure that $\mathbf{x}(r)$ and $\mathbf{u}(r)$ have bounded derivatives as $r \to 0$.

For convenience, we assume that $R = \varnothing$ and that the degenerate constraints are indexed $s+1, s+2, \ldots, m$. Then the left eigenvectors of $J(\mathbf{x}^*, \mathbf{u}^*)$ associated with the eigenvalue zero are $\mathbf{e}_{n+s+1}, \ldots, \mathbf{e}_{n+m}$, where $\mathbf{e}_i$ is the vector with 1 in the ith place and zeros elsewhere. The corresponding bi-orthogonal right eigenvectors will have the form

$$\mathbf{v}_k = \begin{bmatrix} \mathbf{t}_k \\ \alpha_1 \\ \vdots \\ \alpha_s \\ 0 \\ \vdots \\ 0 \\ \beta_k \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad k = s+1, \ldots, m, \tag{2.16}$$

where $\mathbf{t}_k$ must satisfy

$$\nabla g_i(\mathbf{x}^*)\mathbf{t}_k = 0, \quad i = 1, 2, \ldots, s, \tag{2.17}$$

and can be specialized so that

$$\nabla g_k(\mathbf{x}^*)\mathbf{t}_k > 0. \tag{2.18}$$

Substituting in the first n equations of $J(\mathbf{x}^*, \mathbf{u}^*)\, \mathbf{v}_k = 0$ gives

$$\nabla_x^2 L\mathbf{t}_k - \sum_{i=1}^{s} \alpha_i \nabla g_i^{\mathrm{T}} - \beta_k \nabla g_k^{\mathrm{T}} = 0,$$

whence

$$\beta_k = \frac{\mathbf{t}_k^{\mathrm{T}} \nabla_x^2 L\mathbf{t}_k}{\nabla g_k \mathbf{t}_k} > 0, \tag{2.19}$$

using the second-order sufficiency conditions. The condition for (2.10) to have a bounded solution as $r \to 0$ is

$$\mathbf{e}_k^{\mathrm{T}}\left[\frac{d}{dr}(r\phi)\right]_{r=0} = \phi_k(0) = 0, \quad k = s+1, \ldots, m. \tag{2.20}$$

https://doi.org/10.1017/S033427000000240X Published online by Cambridge University Press

In this case the solution is determined only up to arbitrary multplies of the $\mathbf{v}_i$, and can be taken to be of the form

$$\begin{bmatrix} \frac{d\mathbf{x}}{dr} \\ \frac{d\mathbf{u}}{dr} \end{bmatrix} = \begin{bmatrix} \mathbf{w} \\ \mathbf{a} \\ 0 \end{bmatrix} + \sum_{i=s+1}^{m} \theta_i^{(1)} \mathbf{v}_i, \tag{2.21}$$

where $\begin{bmatrix} \mathbf{w} \\ \mathbf{a} \end{bmatrix}$ is determined from the solution trajectory for the nondegenerate problem obtained by deleting the degenerate constraints. This trajectory need not be feasible for the full MPP. To determine the $\theta_i^{(1)}$, $i \in I_2$, it is necessary to ask that the trajectory should have bounded second derivatives when $r=0$. This pattern repeats itself because, in turn, the second derivative values are determined only up to arbitrary multiples of the right eigenvectors, and these multipliers are now found by asking for higher order smoothness.

From (2.10) it follows that the compatibility conditions which must be satisfied at each stage of this process are

$$\frac{d^j}{dr^j}\{u_k(r)\, g_k(\mathbf{x}(r)) - r\phi_k(r)\}\bigg|_{r=0} = 0, \quad k \in I_2, \quad j = 1, 2, \ldots. \tag{2.22}$$

The case $j=1$ just gives (2.20). When $j=2$ we obtain

$$2\frac{du_k}{dr}\ \frac{dg_k}{dr} = 2\frac{d\phi_k}{dr}, \tag{2.23}$$

and when $j=3$

$$3\frac{d^2 u_k}{dr^2}\ \frac{dg_k}{dr} + 3\frac{du_k}{dr}\ \frac{d^2 g_k}{dr^2} = 3\frac{d^2\phi_k}{dr^2}. \tag{2.24}$$

To determine $\theta_i^{(1)}$, $i=s+1, \ldots, m$, we substitute (2.21) into (2.23) to obtain

$$\theta_i^{(1)} \beta_i \left\{ \nabla g_i \left(w + \sum_{j=s+1}^{m} \theta_j^{(1)} \mathbf{t}_j \right) \right\} = \frac{d\phi_i(0)}{dr}, \quad i = s+1, \ldots, m. \tag{2.25}$$

This gives a system of equations for $\theta_j^{(1)}$. If $d\phi_i/dr > 0$, then we require $\theta_i^{(1)} > 0$, as $du_i/dr \geqslant 0$ necessarily. If $d\phi_i/dr = 0$, then either $\theta_i^{(1)} = 0$ or $dg_i/dr = 0$.

If we now attempt to develop $\phi_k(r) = u_k(cr)$ in the same manner as before we find that (2.23) gives either

$$\frac{dg_k}{dr} = \nabla g_k \left\{ \mathbf{w} + \sum_{j=s+1}^{m} \theta_j^{(1)} \mathbf{t}_j \right\} = c \tag{2.26}$$

or

$$\frac{du_k}{dr} = 0, \tag{2.27}$$

https://doi.org/10.1017/S033427000000240X Published online by Cambridge University Press

for each k, $k = s+1, \ldots, m$. Further progress can be made when there is only one degenerate constraint (the case $s = m-1$). Here it is easy to see that if $\mathbf{w}$ corresponds to an infeasible trajectory for the full MPP (so that $\nabla g_m(\mathbf{x}^*)\,\mathbf{w} < 0$) then $\theta_m^{(1)} > 0, c > 0$. On the other hand, if $\mathbf{w}$ is feasible then $u_m(cr) \equiv 0$ provides a solution valid for all $c > 0$, and it appears to be the only one possible for which $(du_m/dr)(0) = 0$. To see this, consider the possibility that $(du_m/dr)(0) = 0$ and $(d^2 u_m/dr^2)(0) > 0$. From (2.24) it follows that we must have

$$\frac{d^2 u_m}{dr^2}(0)\,\nabla g_m(\mathbf{x}^*)\,\mathbf{w} = c^2 \frac{d^2 u_m}{dr^2}(0), \tag{2.28}$$

and $\nabla g_m(\mathbf{x}^*)\mathbf{w}$ is independent of c. This argument extends to higher derivatives by considering the first compatibility condition involving the first nonvanishing derivative of u_m. Thus, in the special case of one degenerate constraint, there appear to be two possible solutions to the differential difference system defining the modified barrier function trajectory, and this raises the question of which (if either) is approached by the solution points generated by the modified algorithm (1.6), (1.7). This does not appear to be an easy question to answer, and we show in the following example that either trajectory may be approached, depending on the particular form of the problem.

EXAMPLE 2.1. To obtain an example in which either of the two possible trajectories is approached by the sequence of solution points, it is convenient to introduce a parameter into the degenerate constraint in Example 1.1. This permits us to vary the feasible region and a result is that both possibilities are important for different ranges of this parameter. The extended problem is:

$$\begin{aligned} &\text{Minimize} && f(\mathbf{x}) = x_2, \\ &\text{subject to} && g_1(\mathbf{x}) = x_2 - x_1^2 \geqslant 0, \\ &\text{and} && g_2(\mathbf{x}) = x_1 + \zeta x_2 \geqslant 0. \end{aligned}$$

The solution is again $f = 0$, and is attained for $\mathbf{x}^* = (0, 0)$ with Lagrange multipliers $\mathbf{u}^* = (1, 0)$, so that g_2 is degenerate for all ζ. We have $J(\mathbf{x}, \mathbf{u})$ given by

$$J(\mathbf{x},\mathbf{u}) = \begin{bmatrix} 2u_1 & 0 & 2x_1 & -1 \\ 0 & 0 & -1 & -\zeta \\ -2x_1 u_1 & u_1 & x_2 - x_1^2 & 0 \\ u_2 & \zeta u_2 & 0 & x_1 + \zeta x_2 \end{bmatrix}.$$

We see that the limiting values for the nondegenerate problem obtained by deleting g_2 are

$$\begin{bmatrix} \mathbf{w} \\ a \end{bmatrix} = \begin{bmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ 0 \end{bmatrix},$$

https://doi.org/10.1017/S033427000000240X Published online by Cambridge University Press

while the right eigenvector for $J(\mathbf{x}^*, \mathbf{u}^*)$ is given by

$$v = \begin{bmatrix} \mathbf{t} \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ -2\zeta \\ 2 \end{bmatrix}.$$

We can now calculate the limiting values for $d\mathbf{x}/dr$ and $d\mathbf{u}/dr$ on the trajectory satisfying $dg_2(\mathbf{x}^*)/dr \doteq c$. We have

$$\nabla g_2\{\mathbf{w}+\gamma\mathbf{t}\} = [1, \zeta]\left\{\begin{bmatrix} 0 \\ 1 \end{bmatrix} + \gamma\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\} = c,$$

giving $\gamma = c - \zeta$, and

$$\begin{bmatrix} \dfrac{d\mathbf{x}}{dr} \\ \dfrac{d\mathbf{u}}{dr} \end{bmatrix} = \begin{bmatrix} c-\zeta \\ 1 \\ -2\zeta(c-\zeta) \\ 2(c-\zeta) \end{bmatrix}.$$

This will define a feasible trajectory for $\zeta \leqslant c$, while $\begin{bmatrix} \mathbf{w} \\ a \end{bmatrix}$ defines a feasible trajectory provided $\zeta > 0$. When $\zeta = c$, the limiting values for the two trajectories become the same. In Fig. 2.1 we sketch the behaviour of $u_2(r)$ for small r, showing the limiting trajectories, and a trajectory for u_2^* defined by minimizing $M(\mathbf{x}, r\mathbf{u}^{(k-1)})$, which has a square root singularity at the origin for each k. This implies an initial fast rise which suggests that, when both trajectories are feasible, the one with the larger value of $(du_2/dr)(x^*)$ will be approached. This is the one with $(du_2/dr)(\mathbf{x}^*) = 2(c-\zeta)$ if $\zeta < c$.

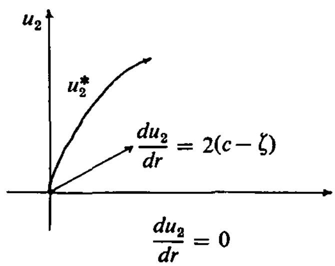

Fig. 2.1. Problem trajectories for small r and $\zeta < c$.

The modified barrier function is

$$M(\mathbf{x}, r_k \mathbf{u}^{(k-1)}) = x_2 - r_k\{u_1^{(k-1)} \log(x_2 - x_1^2) + u_2^{(k-1)} \log(x_1 - \zeta x_2)\},$$

and the conditions for a stationary point give

$$(0, 1) = u_1^{(k)}(-2x_1, 1) + u_2^{(k)}(1, \zeta),$$

https://doi.org/10.1017/S033427000000240X Published online by Cambridge University Press

whence

$$x_1 = \frac{u_2^{(k)}}{2u_1^{(k)}} \quad \text{and} \quad u_1^{(k)} = 1 - \zeta u_2^{(k)},$$

where

$$u_1^{(k)} = \frac{r_k u_1^{(k-1)}}{x_2 - x_1^2} \quad \text{and} \quad u_2^{(k)} = \frac{r_k u_2^{(k-1)}}{x_1 - \zeta x_2}.$$

It follows from the first of these two relations that

$$x_2 = x_1^2 + \frac{r_k u_1^{(k-1)}}{u_1^{(k)}},$$

and we can now express x_1, x_2 and $u_1^{(k)}$ in terms of $u_2^{(k)} = z_k$. Substituting in the expression for x_2 gives

$$z_k \left\{ \frac{z_k}{2(1-\zeta z_k)} + \frac{\zeta}{4} \left(\frac{z_k}{1-\zeta z_k} \right)^2 + \zeta r_k \frac{1-\zeta z_{k-1}}{1-\zeta z_k} \right\} = r_k z_{k-1}.$$

As we are seeking solutions which tend to zero with r, we assume that

$$z_k = \mu_1 r_k + \mu_2 r_k^{1+\sigma} + o(r_k^{1+\sigma}).$$

Substituting for z_k and equating powers of r to zero gives

$$r^2: \quad \tfrac{1}{2}\mu_1^2 + \zeta\mu_1 - c\mu_1 = 0,$$

$$r^{2+\sigma}; \quad \mu_1 \mu_2 + \zeta\mu_2 - \mu_2 c^{1+\sigma} = 0.$$

The first expression vanishes if $\mu_1 = 0$ or $\mu_1 = 2(c-\zeta)$, corresponding to the two trajectories suggested by our analysis, and the second expression permits us to determine σ in each case. We have

(i) $\zeta < c$, $\mu_1 = 2(c-\zeta)$, $\sigma = (\log(2-\zeta/c)/\log c) - 1$;
(ii) $\zeta > c$, $\mu_1 = 0$, $\sigma = (\log \zeta/\log c) - 1$.

This confirms our analysis. To show that both possible limiting trajectories can be approached depending on the value of ζ, we give numerical results for $c = 2$, $\zeta = 1$ and $\zeta = 3$ in Figs. 2.2(a) and 2.2(b). These figures show the possible limiting trajectories corresponding to the two cases $du_2/dr = 2(c-\zeta)$ and $du_2/dr = 0$ as dashed lines, the constraint lines $x_1 + \zeta x_2 = 0$ and $x_2 - x_1^2 = 0$ as broken lines, and a plot of the sequence of values $\{\mathbf{x}(r_k)\}$ as a continuous line. The convergence of this line to the trajectory corresponding to $du_2/dr = 2(c-\zeta)$ when $\zeta = 1$, and to that corresponding to $du_2/dr = 0$ when $\zeta = 3$, is clearly illustrated.

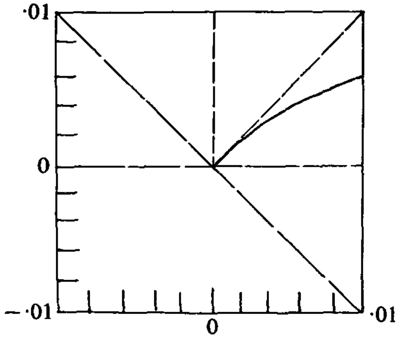

Fig. 2.2(a). ($\xi = 1$, $c = 2$.)

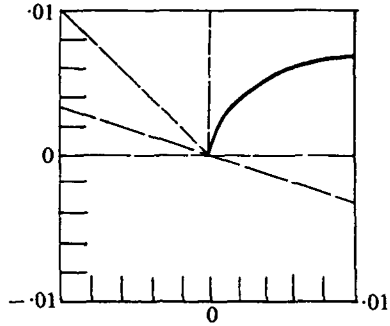

Fig. 2.2(b). ($\xi = 3$, $c = 22$.)

3. Demonstration of $O(r)$ convergence

In this section we establish $O(r)$ convergence of the modified barrier function algorithm (1.6), (1.7). This follows in the nondegenerate case from the super fast convergence to the differentiable trajectory established in Theorem 2.2. Thus the principal interest is in establishing the result in the degenerate case. To do this we are forced to depart from a trajectory analysis and rely on a far from simple proof by contradiction. This is not very satisfying but is perhaps not surprising given the variety of behaviour possible even in the simple example considered in the previous section. The problem lies not only in identifying the particular trajectory to which the successive minima of the modified barrier functions tend, but also in demonstrating that the rate of convergence to this trajectory is fast enough.

To simplify presentation it is convenient to make the following assumptions which can be shown to involve no loss of generality.

(i) All the constraints present are active so that

$$|I_1| + |I_2| = m, \tag{3.1}$$

(ii) $\mathbf{u}^{(0)} = [1, 1, \ldots, 1]$, and

(iii) r_1 is chosen sufficiently small to ensure that $\{\mathbf{x}^{(k)}\}$ belongs to a small enough neighbourhood of $\mathbf{x}^*$. That is, given $\delta \ll 1$, r_1 is chosen small enough such that, for all k,

$$\| \mathbf{x}^{(k)} - \mathbf{x}^* \| \leqslant \delta, \tag{3.2}$$

$$|u_i^{(k)} - u_i^*| \leqslant \delta, \quad i \in I_1, \quad \text{and} \tag{3.3a}$$

$$|u_i^{(k)}| \leqslant \delta, \quad i \in I_2. \tag{3.3b}$$

Before proving the main theorem we derive two preliminary lemmas.

https://doi.org/10.1017/S033427000000240X Published online by Cambridge University Press

LEMMA 3.1. *There exists a bounded constant α such that, for every $k \geqslant 1$,*

$$|u_i^{(k)} - u_i^*| \leqslant \alpha \| \mathbf{x}^{(k)} - \mathbf{x}^* \|, \quad i \in I_1, \tag{3.4a}$$

and

$$|u_i^{(k)}| \leqslant \alpha \| \mathbf{x}^{(k)} - \mathbf{x}^* \|, \quad i \in I_2. \tag{3.4b}$$

PROOF. This result states that the rate of convergence of the multiplier estimates is at least as fast as that of the estimates of the minimizing points. Since $\mathbf{x}^{(k)}$ is a stationary point of $M(\mathbf{x}, r\mathbf{u}^{(k-1)})$, it follows that

$$\nabla f(\mathbf{x}^{(k)}) - \sum_{i \in I} u_i^{(k)} \nabla g_i(\mathbf{x}^{(k)}) = 0. \tag{3.5}$$

Subtracting from (3.5) the Kuhn–Tucker conditions (1.3) gives

$$\nabla_x^2 L(\mathbf{x}^*, \mathbf{u}^*)(\mathbf{x}^{(k)} - \mathbf{x}^*) + o(\| \mathbf{x}^{(k)} - \mathbf{x}^* \|) = \sum_{i \in I_1} (u_i^{(k)} - u_i^*) \nabla g_i(\mathbf{x}^{(k)}) + \sum_{i \in I_2} u_i^{(k)} \nabla g_i(\mathbf{x}^{(k)}). \tag{3.6}$$

Let $\mathbf{w}_i$, $i = 1, \ldots, m$, be the set of vectors of minimum norm bi-orthogonal to the $\nabla g_j(\mathbf{x}^{(k)})$. That such a set exists and is bounded follows, for δ small enough, from the linear independence of the $\nabla g_i(\mathbf{x}^*)$. Then from (3.6) we obtain

$$|u_i^{(k)} - u_i^*| \leqslant \| \mathbf{w}_i^{\mathrm{T}} \nabla_x^2 L(\mathbf{x}^*, \mathbf{u}^*) \| \quad \| \mathbf{x}^{(k)} - \mathbf{x}^* \| + o(\| \mathbf{x}^{(k)} - \mathbf{x}^* \|), \quad i \in I_1, \tag{3.7a}$$

and

$$|u_i^{(k)}| \leqslant \| \mathbf{w}_i^{\mathrm{T}} \nabla_x^2 L(\mathbf{x}^*, \mathbf{u}^*) \| \quad \| \mathbf{x}^{(k)} - \mathbf{x}^* \| + o(\| \mathbf{x}^{(k)} - \mathbf{x}^* \|), \quad i \in I_2. \tag{3.7b}$$

The desired result is an immediate consequence of these inequalities.

LEMMA 3.2. *Let $\kappa_0 > 0$ be chosen sufficiently large. Then there exists $a > 0$ such that, if*

$$r_k \leqslant \frac{1}{\kappa_0} \| \mathbf{x}^{(k)} - \mathbf{x}^* \|, \tag{3.8}$$

then

$$(\mathbf{x}^{(k)} - \mathbf{x}^*)^{\mathrm{T}} \nabla_x^2 L(\mathbf{x}^*, \mathbf{u}^*)(\mathbf{x}^{(k)} - \mathbf{x}^*) \geqslant a \| \mathbf{x}^{(k)} - \mathbf{x}^* \|^2. \tag{3.9}$$

PROOF. We decompose $\mathbf{x}^{(k)} - \mathbf{x}^*$ into

$$\mathbf{x}^{(k)} - \mathbf{x}^* = \| \mathbf{x}^{(k)} - \mathbf{x}^* \| (\mathbf{t}^{(k)} + \mathbf{v}^{(k)}), \tag{3.10}$$

where $\mathbf{v}^{(k)}$ is a linear combination of the $\nabla g_i(\mathbf{x}^*)$, $i \in I_1$, and $\mathbf{t}^{(k)}$ is in the orthogonal complement of this set. The desired result then follows from the second-order sufficiency conditions provided $\| \mathbf{v}^{(k)} \|$ is small enough. For $i \in I_1$ we have

$$g_i(\mathbf{x}^{(k)}) = \frac{r_k u_i^{(k-1)}}{u_i^{(k)}} = r_k(1 + O(\delta))$$

$$\leqslant \frac{(1 + O(\delta))}{\kappa_0} \| \mathbf{x}^{(k)} - \mathbf{x}^* \|, \tag{3.11}$$

so that (3.10) implies that

$$\nabla g_i(\mathbf{x}^*)\,\mathbf{v}^{(k)} = O\left(\max\left(\frac{1+O(\delta)}{\kappa_0}, \|\mathbf{x}^{(k)} - \mathbf{x}^*\|\right)\right), \quad i \in I_1. \tag{3.12}$$

Equation (3.12) is a nonsingular system of equations for the components of $\mathbf{v}^{(k)}$ along each of the $\nabla g_i(\mathbf{x}^*)$, $i \in I_1$. Thus a similar order estimate holds for these components and hence for $\|\mathbf{v}^{(k)}\|$. As $\mathbf{t}^{(k)}$ and $\mathbf{v}^{(k)}$ are orthogonal it follows that

$$\begin{aligned} 1 \geqslant \|\mathbf{t}^{(k)}\| &\geqslant 1 - \|\mathbf{v}^{(k)}\| \\ &\geqslant 1 - O\left(\max\left(\frac{1}{\kappa_0}, \delta\right)\right). \end{aligned} \tag{3.13}$$

Now, applying the second-order sufficiency conditions,

$$\begin{aligned} (\mathbf{x}^{(k)} - \mathbf{x}^*)^{\mathrm{T}} \nabla_x^2 L(\mathbf{x}^*, \mathbf{u}^*)(\mathbf{x}^{(k)} - \mathbf{x}^*) &= \|\mathbf{x}^{(k)} - \mathbf{x}^*\|^2 \{\mathbf{t}^{(k)\mathrm{T}} \nabla_x^2 L(\mathbf{x}^*, \mathbf{u}^*)\mathbf{t}^{(k)} \\ &\qquad + O(\|\mathbf{v}^{(k)}\|)\} \\ &\geqslant a\|\mathbf{x}^{(k)} - \mathbf{x}^*\|^2 \end{aligned}$$

for suitably chosen $a > 0$, provided $\max(1/\kappa_0, \delta)$ is small enough.

THEOREM 3.1. *Let* $\mathbf{x}^*$ *be a regular local solution of the MPP and assume that the problem functions are at least twice continuously differentiable. If* r_1 *is sufficiently small,* $\mathbf{u}^{(0)} > 0$ *and* $c > 1$ *are given, and* $\{\mathbf{x}^{(k)}\}$ *is the sequence of points generated by the modified algorithm, then there exist constants* $0 < \mu_1 < \mu_2$ *such that the asymptotic inequalities*

$$\mu_1 r_k + o(r_k) < \|\mathbf{x}^{(k)} - \mathbf{x}^*\| < \mu_2 r_k + o(r_k) \tag{3.16}$$

hold for sufficiently large k.

PROOF. For $i \in I_1$ we have

$$r_k(1+O(\delta)) = r_k \frac{u_i^{(k-1)}}{u_i^{(k)}} = g_i(\mathbf{x}^{(k)}) = \nabla g_i(\mathbf{x}^*)(\mathbf{x}^{(k)} - \mathbf{x}^*) + o(\|\mathbf{x}^{(k)} - \mathbf{x}^*\|), \tag{3.17}$$

so that, for k sufficiently large, there exists K_1 such that

$$r_k \leqslant K_1 \|\mathbf{x}^{(k)} - \mathbf{x}^*\|.$$

Thus it is necessary only to establish the right-hand inequality in (3.16). Assume this is false. Then there exists a subsequence $\{k(l)\}$ such that

$$r_{k(l)} < \frac{1}{\kappa_l}\|\mathbf{x}^{(k(l))} - \mathbf{x}^*\|, \tag{3.18}$$

with $\{\kappa_l\} \to \infty$. Choose l sufficiently large so that (3.8) is satisfied. Multiplying (3.6) by $(\mathbf{x}^{(k)} - \mathbf{x}^*)$ and using that

$$g_i(\mathbf{x}^{(k)}) = \nabla g_i(\mathbf{x}^*)(\mathbf{x}^{(k)} - \mathbf{x}^*) + o(\|\mathbf{x}^{(k)} - \mathbf{x}^*\|), \quad i \in I,$$

L

it follows that

$$(\mathbf{x}^{(k)}-\mathbf{x}^*)^T \nabla_x^2 L(\mathbf{x}^*,\mathbf{u}^*)(\mathbf{x}^{(k)}-\mathbf{x}^*)+o(\|\mathbf{x}^{(k)}-\mathbf{x}^*\|^2)-\sum_{i\in I_1}(u_i^{(k)}-u_1^*)(g_i(\mathbf{x}^{(k)}) + o(\|\mathbf{x}^{(k)}-\mathbf{x}^*\|)) = \sum_{i\in I_2}(u_i^{(k)}(g_i\mathbf{x}^{(k)})+o(\|\mathbf{x}^{(k)}-\mathbf{x}^*\|)).$$

By further adjusting the constant a to allow for terms of smaller order, it follows from (3.9) that, for $k=k(l)$ and l large enough,

$$a\|\mathbf{x}^{(k)}-\mathbf{x}^*\|^2-\sum_{i\in I_1}|u_i^{(k)}-u_i^*|g_i(\mathbf{x}^{(k)}) \leqslant \sum_{i\in I_2} u_i^{(k)} g_i(\mathbf{x}^{(k)}). \tag{3.19}$$

Using (3.4a), (3.17) and (3.18), for $i\in I_1$ we have

$$|u_i^{(k)}-u_i^*|g_i(\mathbf{x}^{(k)}) \leqslant \frac{\alpha(1+O(\delta))\|\mathbf{x}^{(k)}-\mathbf{x}^*\|^2}{\kappa_l}. \tag{3.20}$$

Thus the first term on the left-hand side of (3.19) dominates the second, and we can write

$$\begin{aligned}\tfrac{1}{2}a\|\mathbf{x}^{(k)}-\mathbf{x}^*\|^2 &\leqslant \sum_{i\in I_2} u_i^{(k)} g_i(\mathbf{x}^{(k)})\\ &= \sum_{i\in I_2} r_k u_i^{(k-1)}\\ &\leqslant r_k|I_2|\alpha\|\mathbf{x}^{(k-1)}-\mathbf{x}^*\|.\end{aligned} \tag{3.21}$$

By (3.18),

$$\tfrac{1}{2}a\kappa_l^2 r_k^2 \leqslant r_k|I_2|\alpha\|\mathbf{x}^{(k-1)}-\mathbf{x}^*\|,$$

so that

$$r_{k-1} \leqslant \frac{2c\alpha|I_2|}{a\kappa_l^2}\|\mathbf{x}^{(k-1)}-\mathbf{x}^*\|. \tag{3.22}$$

Thus condition (3.8) of Lemma 3.2 is satisfied for $k:=k-1$, and the above arguments can be repeated using (3.22) instead of (3.18). But then, by back-tracking,

$$\|\mathbf{x}^{(1)}-\mathbf{x}^*\| \geqslant \left\{\frac{a\kappa_l}{2c\alpha|I_2|}\right\}^{\sum_{i=0}^{k-2} 2^i} \kappa_l r_1. \tag{3.25}$$

This contradicts (3.2) because κ_l can be chosen arbitrarily large.

4. Numerical results

In this section numerical results are presented for the use of extrapolation in conjunction with the modified barrier function algorithm, and justification for the development of the new algorithm is obtained by comparing its numerical performance with that of the extrapolated classical barrier function algorithm discussed in our previous paper [4]. We note that we must expect important differences in the use of extrapolation between degenerate and nondegenerate problems for, while it can be justified to high order for nondegenerate problems as a result of the super fast convergence to the smooth trajectory defined by (2.4), in the degenerate case we have been able to show only $O(r)$ convergence and examples 1.1 and 2.1 suggest that further expansion in integral powers of r is not possible in general. However, in practice it may not be easy to distinguish between degenerate problems and nondegenerate problems with small multipliers (these are the ill-conditioned problems considered in [4]), and for this reason we elect to build up the table of extrapolations and to estimate from this the order of convergence supported. For example, in degenerate cases the first extrapolated column should show an improvement in its rate of convergence and this should not be greatly changed in succeeding columns provided c is chosen to ensure the extrapolation is stable.

In implementing the modified algorithm the key quantities that have to be specified are c, r_1 and $\mathbf{u}^{(0)}$. Here we have chosen $c = 2$. This choice is known to be stable for extrapolation (Laurent [5]) and represents a compromise between having c too small with the associated danger of instability in the extrapolation, and having c too large so that only a few steps of the algorithm are possible before the small values of r makes the successive minimizations increasingly difficult. To determine an initial value of $\mathbf{u}$, a preliminary minimization is carred out with $\mathbf{u}^T = [1, 1, \ldots, 1]$ and the result is used to determine $\mathbf{u}^{(0)}$. Choice of the value of $\bar{r}$ to use in this minimization must depend on the scale of the problem. If $\bar{r}$ is chosen too small then the resulting minimization will be difficult, but it is important that $\bar{r}$ be chosen small enough to ensure that our local problem analysis is valid. We have found a rule of the form

$$\bar{r} = \min\left\{10^{-2}, \frac{\|\nabla f\| 10^{-2}}{n}\right\}$$

to be satisfactory. The dependence on $\|\nabla f\|$ seems reasonable as this is linked through the necessary conditions for a minimum to the size of the $1/g_i$, $i = 1, 2, \ldots, m$, and the size of these quantities, in turn, reflects the difficulty in minimizing $B(\mathbf{x}, \bar{r})$. The value of r_1 is now taken as

$$r_1 = \min\left\{\frac{10^{-2}}{\|\mathbf{u}^{(0)}\|_\infty}, \frac{\|\nabla f(\mathbf{x}(\bar{r}))\| 10^{-2}}{n\|\mathbf{u}^{(0)}\|_\infty}\right\},$$

where the difference between this expression and that for $\bar{r}$ takes account of the inclusion of the Lagrange multipliers as weights in the modified algorithm.

https://doi.org/10.1017/S033427000000240X Published online by Cambridge University Press

In order to have a meaningful comparison between the modified and classical algorithms we must make the first iterations compatible. Let p be the controlling parameter in the classical algorithm. Then we select $p_1 = r_1 \| \mathbf{u}^{(0)} \|_\infty$ and then define $p_{k+1} = p_k/2$, $k = 1, 2, \ldots$. The numerical performance of the two algorithms is examined and compared in terms of the quantities $\mathrm{LB}(\mathbf{x})$, $\mathrm{UB}(\mathbf{x})$ which specify lower and upper bounds for the values of the constraints $g_i(\mathbf{x})$, $i \in I$. We have

$$\mathrm{LB}(x) = \min_{i \in I} | g_i(\mathbf{x}) |,$$

$$\mathrm{UB}(\mathbf{x}) = \max_{i \in I} | g_i(\mathbf{x}) |.$$

The quantity $\mathrm{LB}(\mathbf{x}^{(k)})$ shows the approach of $x^{(k)}$ to the constraints and hence indicates the degree of difficulty associated with the unconstrained minimization. The numerical results have shown that our choice of p_1 is such that the quantities $\mathrm{LB}(\mathbf{x}^{(k)})$ have approximately the same values both for the modified and classical algorithm. The quantity $\mathrm{UB}(\mathbf{x}^{(k)})$ shows the convergence of $\mathbf{x}^{(k)}$ to $\mathbf{x}^*$. For the classical algorithm applied to nondegenerate problems the ratio $\mathrm{UB}(\mathbf{x}^{(k)})/\mathrm{LB}(\mathbf{x}^{(k)})$ is asymptotically proportional to γ, the measure of degeneracy, and it is unbounded with $(c)^{k/2}$ for degenerate problems. However, for the modified algorithm it follows from (2.15) that the ratio tends to 1 in the nondegenerate case, while in the degenerate case it follows from (3.16) and (3.17) that it tends to a constant.

Test problems have been chosen

(a) to compare the performance of the modified and classical algorithms on nondegenerate problems, and

(b) to exemplify the behaviour of the modified algorithm on a range of degenerate problems.

As a basis for our comparison between the modified and classical algorithm, we have used the test problems considered in [3]. We give results for the easy Rosen–Suzuki problem, Colville problem I and the more difficult Colville problem II. The degenerate problems considered include Example 1.1, a modified form of the Rosen–Suzuki problem, and a highly degenerate problem suggested to us by Professor J. B. Rosen. These last two problems are detailed in Appendix 1.

The numerical results are summarized in Tables 4.1 to 4.6. They are expressed in terms of the quantities $\mathrm{LB}(\mathbf{x}^{(k)})$, $\mathrm{UB}(\mathbf{x}^{(k)})$ and $\mathrm{UB}(\mathbf{x}^{(k)}_{\min})$, where $\mathbf{x}^{(k)}_{\min}$ is the extrapolated value at the kth stage.

(i) *Nondegenerate problems*. For nondegenerate problems with reasonably large values of γ the modified algorithm should prove better than the classical algorithm. This is verified by the results listed in Tables 4.2 and 4.3 for the two Colville problems, while the results in Table 4.1 for the Rosen–Suzuki problem show the classical algorithm performing better than the modified algorithm. This behaviour is expected as $\gamma = 1.6$ and hence the results from the classical algorithm extrapolate well. On the other hand, the first few

https://doi.org/10.1017/S033427000000240X Published online by Cambridge University Press

iterations of the modified algorithm are needed for the convergence of $\{\mathbf{x}^{(k)}\}$ to the trajectory $\{\hat{\mathbf{x}}(r_k)\}$. However, in all cases the performance of the modified algorithm is effectively independent of γ, in marked contrast with the performance of the classical algorithm. In this important respect the performance of the modified algorithm is superior to that of the classical algorithm.

(ii) *Degenerate problems.* For degenerate problems the results given in Tables 4.4 to 4.6 show clearly that the modified algorithm is superior to the classical algorithm. The classical algorithm performs especially badly on the Rosen-Kreuser problem as detailed in Table 4.6. Note also that the improvement from extrapolation is only first order in the second and third examples. There is some tendency for the extrapolated quantities to follow an $O(r^{\frac{3}{2}})$ behaviour, but this is certainly not nearly so evident as the $O(r)$ behaviour before extrapolation. Example 1.1 provides an exception but there the higher order improvement is possible only when $c=2$.

We remark that the ability of the modified method to drive out the inactive constraints was clearly shown in our calculations. For while the multiplier estimates for these constraints produced by the classical method gave values of $O(10^{-4})$, the corresponding values for the modified method were $O(10^{-24})$.

The numerical results were produced on the Univac 1100/42 at the Australian National University. The minimizations were carried out using the unconstrained minimization subroutine FUNMIN with the tolerance adjusted so that at least 9 correct figures were obtained in $\mathbf{x}$.

TABLE 4.1

Numerical results for the Rosen–Suzuki problem ($\gamma=1.6$, $p_1=10^{-2}$)

	Classical algorithm			Modified algorithm		
k	$\mathrm{LB}(\mathbf{x}^{(k)})$	$\mathrm{UB}(\mathbf{x}^{(k)})$	$\mathrm{UB}(\mathbf{x}^{(k)}_{\min})$	$\mathrm{LB}(\mathbf{x}^{(k)})$	$\mathrm{UB}(\mathbf{x}^{(k)})$	$\mathrm{UB}(\mathbf{x}^{(k)}_{\min})$
1	0.50×10^{-2}	0.10×10^{-1}	0.10×10^{-1}	0.50×10^{-2}	0.50×10^{-2}	0.50×10^{-2}
2	0.25×10^{-2}	0.50×10^{-2}	0.93×10^{-4}	0.25×10^{-2}	0.25×10^{-2}	0.67×10^{-4}
3	0.12×10^{-2}	0.25×10^{-2}	0.73×10^{-6}	0.12×10^{-2}	0.12×10^{-2}	0.20×10^{-4}
4	0.62×10^{-3}	0.13×10^{-2}	0.10×10^{-8}	0.62×10^{-3}	0.62×10^{-3}	0.27×10^{-5}
5	0.31×10^{-3}	0.63×10^{-3}	0.14×10^{-9}	0.31×10^{-3}	0.31×10^{-3}	0.16×10^{-6}
6				0.16×10^{-3}	0.16×10^{-3}	0.40×10^{-8}
7				0.77×10^{-3}	0.77×10^{-3}	0.41×10^{-8}

TABLE 4.2
Numerical results for the Colville problem I ($\gamma=91$, $p_1=10^{-2}$)

	Classical algorithm			Modified algorithm		
k	$LB(\mathbf{x}^{(k)})$	$UB(\mathbf{x}_{min}^{(k)})$	$UB(\mathbf{x}^{(k)})$	$LB(\mathbf{x}^{(k)})$	$UB(\mathbf{x}^{(k)})$	$UB(\mathbf{x}_{min}^{(k)})$
1	0.84×10^{-3}	0.66×10^{-1}	0.66×10^{-1}	0.84×10^{-3}	0.12×10^{-2}	0.12×10^{-2}
2	0.42×10^{-3}	0.38×10^{-1}	0.10×10^{-1}	0.42×10^{-3}	0.42×10^{-3}	0.35×10^{-3}
3	0.21×10^{-3}	0.21×10^{-1}	0.19×10^{-2}	0.21×10^{-3}	0.21×10^{-3}	0.11×10^{-3}
4	0.11×10^{-3}	0.11×10^{-1}	0.29×10^{-3}	0.10×10^{-3}	0.10×10^{-3}	0.15×10^{-4}
5	0.53×10^{-4}	0.58×10^{-2}	0.23×10^{-4}	0.52×10^{-4}	0.52×10^{-4}	0.98×10^{-6}
6	0.26×10^{-4}	0.29×10^{-2}	0.16×10^{-5}	0.26×10^{-4}	0.26×10^{-4}	0.28×10^{-7}
7	0.13×10^{-4}	0.15×10^{-2}	0.52×10^{-7}	0.13×10^{-4}	0.13×10^{-4}	0.24×10^{-9}
8	0.66×10^{-4}	0.75×10^{-3}	0.14×10^{-8}			
	$\lVert x_{min}^{(8)}-x_{min}^{(7)}\rVert\leqslant10^{-10}$					

TABLE 4.3
Numerical results for the Colville problem II ($\gamma=83$, $p_1=10^{-2}$)

	Classical algorithm			Modified algorithm		
k	$LB(\mathbf{x}^{(k)})$	$UB(\mathbf{x}^{(k)})$	$UB(\mathbf{x}_{min}^{(k)})$	$LB(\mathbf{x}^{(k)})$	$UB(\mathbf{x}^{(k)})$	$UB(\mathbf{x}_{min}^{(k)})$
1	0.18×10^{-3}	0.43×10^{-1}	0.43×10^{-1}	0.18×10^{-3}	0.18×10^{-3}	0.18×10^{-3}
2	0.88×10^{-4}	0.21×10^{-1}	0.11×10^{-2}	0.88×10^{-4}	0.88×10^{-4}	0.80×10^{-5}
3	0.44×10^{-4}	0.11×10^{-1}	0.12×10^{-3}	0.44×10^{-4}	0.44×10^{-4}	0.27×10^{-5}
4	0.22×10^{-4}	0.55×10^{-2}	0.15×10^{-4}	0.22×10^{-4}	0.22×10^{-4}	0.38×10^{-6}
5	0.11×10^{-4}	0.28×10^{-2}	0.13×10^{-5}	0.11×10^{-5}	0.11×10^{-4}	0.25×10^{-7}
6	0.55×10^{-5}	0.14×10^{-2}	0.73×10^{-7}	0.55×10^{-5}	0.55×10^{-5}	0.86×10^{-9}
7	0.28×10^{-5}	0.70×10^{-3}	0.22×10^{-8}			
8	0.13×10^{-5}	0.35×10^{-3}	0.14×10^{-9}			

TABLE 4.4
Numerical results for Example 1.1 ($p_1=0.50\times10^{-2}$)

	Classical algorithm			Modified algorithm		
k	$LB(\mathbf{x}^{(k)})$	$UB(\mathbf{x}^{(k)})$	$UB(\mathbf{x}_{min}^{(k)})$	$LB(\mathbf{x}^{(k)})$	$UB(\mathbf{x}^{(k)})$	$UB(\mathbf{x}_{min}^{(k)})$
1	0.25×10^{-2}	0.35×10^{-1}	0.35×10^{-1}	0.25×10^{-2}	0.94×10^{-2}	0.94×10^{-2}
2	0.13×10^{-2}	0.25×10^{-1}	0.15×10^{-1}	0.13×10^{-2}	0.34×10^{-2}	0.25×10^{-2}
3	0.63×10^{-3}	0.18×10^{-1}	0.89×10^{-2}	0.63×10^{-3}	0.15×10^{-2}	0.18×10^{-3}
4	0.31×10^{-3}	0.12×10^{-1}	0.59×10^{-2}	0.31×10^{-3}	0.68×10^{-3}	0.44×10^{-5}
5	0.16×10^{-3}	0.88×10^{-2}	0.41×10^{-2}	0.16×10^{-3}	0.33×10^{-3}	0.41×10^{-7}
6	0.78×10^{-4}	0.62×10^{-2}	0.28×10^{-2}	0.78×10^{-4}	0.16×10^{-3}	0.16×10^{-9}
7	0.39×10^{-5}	0.44×10^{-2}	0.20×10^{-2}			
8	0.20×10^{-5}	0.31×10^{-2}	0.14×10^{-2}			
9	0.98×10^{-5}	0.22×10^{-2}	0.99×10^{-3}			
10	0.49×10^{-5}	0.16×10^{-2}	0.70×10^{-3}			
11	0.24×10^{-5}	0.11×10^{-2}	0.50×10^{-3}			
	not converge after 11 iterations					

TABLE 4.5
Numerical results for the modified Rosen–Suzuki problem ($p_1 = 10^{-2}$)

	Classical algorithm			Modified algorithm		
k	$LB(\mathbf{x}^{(k)})$	$UB(\mathbf{x}^{(k)})$	$UB(\mathbf{x}^{(k)}_{\min})$	$LB(\mathbf{x}^{(k)})$	$UB(\mathbf{x}^{(k)})$	$UB)\mathbf{x}^{(k)}_{\min})$
1	0.48×10^{-2}	0.13	0.13	0.41×10^{-2}	0.24×10^{-1}	0.24×10^{-1}
2	0.24×10^{-2}	0.88×10^{-1}	0.49×10^{-1}	0.23×10^{-2}	0.81×10^{-2}	0.80×10^{-2}
3	0.21×10^{-2}	0.62×10^{-1}	0.30×10^{-1}	0.12×10^{-2}	0.32×10^{-2}	0.49×10^{-3}
4	0.61×10^{-3}	0.43×10^{-1}	0.20×10^{-1}	0.59×10^{-3}	0.14×10^{-2}	0.39×10^{-4}
5	0.31×10^{-3}	0.30×10^{-1}	0.14×10^{-1}	0.29×10^{-3}	0.66×10^{-3}	0.86×10^{-5}
6	0.15×10^{-3}	0.21×10^{-1}	0.97×10^{-2}	0.14×10^{-3}	0.32×10^{-3}	0.28×10^{-5}
7	0.78×10^{-4}	0.15×10^{-1}	0.68×10^{-2}	0.74×10^{-4}	0.15×10^{-3}	0.87×10^{-6}
8	0.39×10^{-4}	0.11×10^{-1}	0.48×10^{-2}	0.37×10^{-4}	0.76×10^{-4}	0.25×10^{-6}
9	0.19×10^{-4}	0.75×10^{-2}	0.34×10^{-2}	0.19×10^{-4}	0.38×10^{-4}	0.80×10^{-7}
10	0.97×10^{-5}	0.53×10^{-2}	0.24×10^{-2}	0.93×10^{-5}	0.19×10^{-4}	0.27×10^{-7}
11	0.49×10^{-5}	0.38×10^{-2}	0.17×10^{-2}	$\lVert\mathbf{x}^{(10)}_{\min}-\mathbf{x}^{(9)}_{\min}\rVert\leqslant10^{-10}$, the estimated solution, $\mathbf{x}^{(10)}_{\min}$, is corrected to 9 decimal places		
	not converge after 11 iterations					

TABLE 4.6
Numerical results for the Rosen–Kreuser problem ($p_1 = 10^{-2}$)

	Classical algorithm			Modified algorithm		
k	$LB(\mathbf{x}^{(k)})$	$UB(\mathbf{x}^{(k)})$	$UB(\mathbf{x}^{(k)}_{\min})$	$LB(\mathbf{x}^{(k)})$	$UB(\mathbf{x}^{(k)})$	$UB(\mathbf{x}^{(k)}_{\min})$
1	0.50×10^{-2}	15.9	15.9	0.50×10^{-2}	0.34	0.34
2	0.25×10^{-2}	11.3	6.67	0.25×10^{-2}	0.36×10^{-1}	0.27
3	0.13×10^{-2}	8.01	4.08	0.17×10^{-2}	0.85×10^{-2}	0.65×10^{-1}
4	0.63×10^{-3}	5.69	2.71	0.63×10^{-3}	0.28×10^{-2}	0.60×10^{-2}
5	0.31×10^{-3}	4.03	1.87	0.31×10^{-3}	0.12×10^{-2}	0.18×10^{-3}
6	0.16×10^{-3}	2.85	1.30	0.15×10^{-3}	0.52×10^{-3}	0.13×10^{-4}
7	0.78×10^{-4}	2.02	0.92	0.78×10^{-4}	0.24×10^{-3}	0.37×10^{-5}
8	0.39×10^{-4}	1.43	0.65	0.39×10^{-4}	0.12×10^{-3}	0.10×10^{-5}
9	0.20×10^{-4}	1.01	0.46	0.20×10^{-4}	0.57×10^{-4}	0.37×10^{-6}
10	0.98×10^{-5}	0.72	0.32	$\lVert\mathbf{x}^{(9)}_{\min}-\mathbf{x}^{(8)}_{\min}\rVert\leqslant10^{-10}$, the estimated solution, $\mathbf{x}^{(9)}_{\min}$, is corrected to 9 decimal places		
11	0.49×10^{-5}	0.51	0.23			
	not converge after 11 iterations					

Appendix

The Rosen–Suzuki problem. (Rosen and Suzuki (1965).) $n=4$, $m=3$.

$$
\begin{aligned}
\text{Minimize}\quad & f(\mathbf{x}) = x_1^2+x_2^2+2x_3^2+x_4^2-5x_1-5x_2-21x_3+7x_4,\\
\text{subject to}\quad & g_1(\mathbf{x}) = -x_1^2-x_2^2-x_3^2-x_4^2-x^2-x_1+x_2-x_3+x_4+8 \geqslant 0,\\
& g_2(\mathbf{x}) = -x_1^2-2x_2^2-x_3^2-2x_4^2+x_1+x_4+10 \geqslant 0,\\
\text{and}\quad & g_3(\mathbf{x}) = -2x_1^2-x_2^2-x_3^2-2x_1+x_2+x_4+5 \geqslant 0.
\end{aligned}
$$

Constraints 1 and 3 are active. The constrained minimum is at $\mathbf{x}^* = (0,1,2,-1)$. A feasible starting point is $\mathbf{x}^{(0)} = 0$.

https://doi.org/10.1017/S033427000000240X Published online by Cambridge University Press

The modified Rosen–Suzuki problem. $n = 4$, $m = 3$.

This problem is obtained by subtracting 1 from the second constraint of the Rosen–Suzuki problem. The (modified) second constraint,

$$g_2(\mathbf{x}) = -x_1^2 - 2x_2^2 - x_3^2 - 2x_4^2 + x_1 + x_4 + 9 \geqslant 0,$$

is degenerate at the solution.

The Rosen–Kreuser problem. $n = 15$, $m = 10$.

$$\text{Minimize} \quad f(\mathbf{x}) = -\sum_{i=1}^{15} c_i x_i,$$

$$\text{subject to} \quad g_i(\mathbf{x}) = b_i - \sum_{j=1}^{15} a_{ij} x_j^2 \geqslant 0, \quad i = 1, 2, \ldots, 10,$$

where the coefficients a_{ij}, b_i and c_i are given in Table A.1. All constraints are active and the first nine constraints are degenerate. The constrained minimum is at $\mathbf{x}^* = (1, 1, \ldots, 1)$.

TABLE A.1
Coefficients for the Rosen–Kreuser problem

i \ j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
c_j	20	40	400	20	80	20	40	140	380	280	80	40	140	40	120	b_j
1	100	100	10	5	10	0	0	25	0	10	55	5	45	20	0	385
2	90	100	10	35	20	5	0	35	55	25	20	0	40	25	10	470
3	70	50	0	55	25	100	40	50	0	30	60	10	30	0	40	560
4	50	0	0	65	35	100	35	60	0	15	0	75	35	30	65	565
a_{ij}																
5	50	10	70	60	45	45	0	35	65	5	75	100	75	10	0	645
6	40	0	50	95	50	35	10	60	0	45	15	20	0	5	5	430
7	30	60	30	90	0	30	5	25	0	70	20	25	70	15	15	485
8	20	30	40	25	40	25	15	10	80	20	30	30	5	65	20	455
9	10	70	10	35	25	65	0	30	0	0	25	0	15	50	55	390
10	5	10	100	5	20	5	10	35	95	70	20	10	35	10	30	460

References

[1] A. V. Fiacco and G. P. McCormick, *Nonlinear programming: sequential unconstrained minimization techniques* (Wiley, 1968).

[2] K. Jittorntrum, "Sequential algorithms in nonlinear programming", (Ph.D. thesis, Australian National University, 1978).

[3] K. Jittorntrum, "Accelerated convergence for the Powell/Hestenes multiplier method", *Math. Programming* (1979) (to appear).

[4] K. Jittorntrum and M. R. Osborne, "Trajectory analysis and extrapolation in barrier function methods", *J. Austral. Math. Soc. B* 21 (1978), 1–18.

[5] P. J. Laurent, "Convergence de procédé d'extrapolation de Richardson", *Troisième Congrès de l'AFCALTI* (Toulouse, 1963), 81–98.

[6] M. R. Osborne, "Topics in optimization" (Comp. Sci. Dept., Stanford University, STAN–CS–72–279).

[7] M. J. D. Powell, "A fast algorithm for nonlinearly constrained optimization calculations", in *Numerical analysis* (Lecture notes in mathematics No. 630, New York: Springer-Verlag, 1978), 144–157.

Department of Mathematics
Chiengmai University
Cheingmai
Thailand

Department of Statistics
Institute of Advanced Studies
Australian National University
Canberra
A.C.T. 2600