J. Austral. Math. Soc. **20** (Series B) (1978), 352–369

TRAJECTORY ANALYSIS AND EXTRAPOLATION IN BARRIER FUNCTION METHODS

KRISORN JITTORNTRUM and M. R. OSBORNE

(Received 21 February 1978)

Abstract

It has been known for some time that if a certain non-degeneracy condition is satisfied then the successive solution estimates $\mathbf{x}(r)$ produced by barrier function techniques lie on a smooth trajectory. Accordingly, extrapolation methods can be used to calculate $\mathbf{x}(0)$. In this paper we analyse the situation further treating the special case of the log barrier function. If the non-degeneracy assumption is not satisfied then the approach to $\mathbf{x}(0)$ is like $r^{\frac{1}{2}}$ rather than like r which would be expected in the non-degenerate case. A measure of sensitivity is introduced which becomes large when the non-degeneracy assumption is close to violation, and it is shown that this sensitivity measure is related to the growth of $d^i\mathbf{x}/dr^i$ with respect to i for fixed r small enough on the solution trajectory. With this information it is possible to analyse the extrapolation procedure and to predict the number of stages of extrapolation which are useful.

1. Introduction

In this paper we consider the solution of the mathematical programming problem (MPP)

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

where f and the g_i, $i = 1, 2, \ldots, m$, are appropriately smooth functions on $\mathrm{R}^n \to R$, by means of the sequential minimization of the 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 $r_1 > r_2 > \ldots > r_k > \ldots$ where $\lim_{k\to\infty} r_k = 0$. Let $\mathbf{x}(r_k)$ be the exact minimum of (1.2) produced by some algorithmic procedure for $r = r_k$. We assume that this minimum exists and is well defined. It is well known that the limit points

of the sequence $\{\mathbf{x}(r_k)\}$ are solutions to the MPP under very general conditions [2]. However, to ensure that the solution points lie on a smooth trajectory parameterized by r we require the problem to have considerably more structure. Proofs of the following results can be found in [2].

PROPOSITION (*first-order necessary conditions*). *A necessary condition for* $\mathbf{x}^* \in S$ *to be a solution of the MMP is that there exist multipliers* u_i, $i = 1, 2, \ldots, m$, *satisfying the Kuhn–Tucker conditions*

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

and

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

and that the constraint set S *satisfy a suitable regularity condition at* $\mathbf{x}^*$.

The Lagrangian function for the MMP is

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

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

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

for all $\mathbf{t}$ *such that for each* i *for which* $u_i > 0$

$$\nabla g_i(\mathbf{x}^*) \mathbf{t} = 0$$

then there exists an open neighbourhood N *of* $\mathbf{x}^*$ *in* S *such that if* $\mathbf{x} \in N$, $\mathbf{x} \neq \mathbf{x}^*$, *then* $f(\mathbf{x}) > f(\mathbf{x}^*)$. *The second-order sufficiency conditions ensure that* $\mathbf{x}^*$ *is an* isolated *solution of the MPP.*

REMARK 1.1. Clearly there exists $M > 0$ such that

$$\mathbf{t}^{\mathrm{T}} \nabla_x^2 L(\mathbf{x}^*, \mathbf{u}) \mathbf{t} \leqslant M \|\mathbf{t}\|^2 \tag{1.6}$$

for all $\mathbf{t}$ satisfying the requirements of the second-order sufficiency conditions.

It is convenient for our purposes to classify the constraints into several sets depending on their behaviour at $\mathbf{x}^*$. Let

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

https://doi.org/10.1017/S0334270000001715 Published online by Cambridge University Press

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

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

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

DEFINITION. The MMP is *non-degenerate* if $I_2 = \emptyset$.

DEFINITION. $\mathbf{x}^*$ is a *regular local solution* of the MPP if
(i) the sequence $\{\mathbf{x}(r_k)\} \to \mathbf{x}^*$,
(ii) $\nabla g_i(\mathbf{x}^*)$, $i \in I$, are linearly independent, and
(iii) the second-order sufficiency conditions hold at $\mathbf{x}^*$.

REMARK 1.2. 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^* 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.

In what follows it will be convenient to denote $r_k / g_i(\mathbf{x}(r_k))$ by $u_i(r_k)$, and $A(r_k) - A(0)$ by $\Delta_k(A)$ for any function $A(r)$.

DEFINITION. The Jacobian of the system

$$\nabla f(\mathbf{x}(r)) - \sum_{i=1}^{m} u_i(r) \nabla g_i(\mathbf{x}(r)) = 0,$$

$$u_i(r) g_i(\mathbf{x}(r)) = r, \quad i = 1, 2, \ldots, m,$$

with respect to $\mathbf{x}$, $\mathbf{u}$ is given by

$$J(r) = \begin{bmatrix} \nabla_x^2 L(\mathbf{x}(r), \mathbf{u}(r)) & \cdots & -\nabla g_i(\mathbf{x}(r))^{\mathrm{T}} & \cdots \\ \vdots & & & \\ u_i(r) \nabla g_i(\mathbf{x}(r)) & & g_i(\mathbf{x}(r)) & \\ \vdots & & & \ddots \end{bmatrix}. \tag{1.7}$$

It is called the *Jacobian* of the MPP.

https://doi.org/10.1017/S0334270000001715 Published online by Cambridge University Press

THEOREM 1.1. *Let* $\mathbf{x}^*$ *be a regular local solution of a non-degenerate MPP, then*

(i) $J(0)$ *is non-singular,*

(ii) $\mathbf{x}(r)$, $\mathbf{u}(r)$ *lie on a smooth trajectory uniquely determined for* r *small enough by the system of differential equations*

$$J(r)\begin{bmatrix} d\mathbf{x}/dr \\ d\mathbf{u}/dr \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{e} \end{bmatrix} \tag{1.8}$$

and the initial conditions $\mathbf{x}(0) = \mathbf{x}^*$, $\mathbf{u}(0) = \mathbf{u}^*$, *where* $\mathbf{e}$ *is the vector of length* m *each component of which is* 1, *and*

(iii) $\|\Delta_k(\mathbf{x})\|, \|\Delta_k(\mathbf{u})\| = O(r_k)$.

Theorem 1.1 provides the basic information necessary to use standard extrapolation procedures to obtain an estimate for $\mathbf{x}^*$, $\mathbf{u}^*$ from the results of minimizing $B(\mathbf{x}, r)$ for (say) $r = r_1, r_2, \ldots, r_k$. Standard linear extrapolation applied to $x_i(r)$ on the points $r_1, r_2, \ldots, r_k$ is equivalent to first fitting a Lagrange interpolation polynomial of degree $k-1$ to the data say $P_{k-1}(\{x_i(r_j), j = 1, \ldots, k\}; r)$, and then evaluating this at $r = 0$. A standard argument gives that the error in this estimate is

$$x_i^* - P_{k-1}(\{x_i(r_j), j = 1, \ldots, k\}; 0) = \frac{(-1)^{k+1}}{k!} \prod_{j=1}^{k} r_j \frac{d^k x_i}{dr^k}(\xi), \tag{1.9}$$

where ξ is a mean value. The extrapolation is known to be numerically stable provided r_{j+1}/r_j is small enough for each j [5].

The numerical performance of extrapolation when used to improve the basic barrier function algorithm has been considered by several authors (for example, [2, 4, 6]). Although the resulting algorithms tend to be very robust, the improvement in efficiency is not always satisfactory. Our intention here is to examine the implication of the non-degeneracy assumption $I_2 = \emptyset$, and in the next section we show that the trajectory analysis summarized in Theorem 1.1 must be modified if this condition is relaxed. It proves to be possible to introduce a "measure of non-degeneracy" and this is done in Section 3 where it is also shown that this measure is an important parameter in determining the growth of $d^i\mathbf{x}/dr^i$, $d^i\mathbf{u}/dr^i$ as i increases at a point on the solution trajectory. In Section 4 a possible strategy for estimating the worth of successive extrapolation steps is suggested and illustrated by numerical results.

It is not claimed that the results presented here provide a basis for re-examining the value of the basic sequential minimization plus extrapolation procedures. However, we believe that they do improve our understanding of the problems caused by degeneracy, and this in itself should be useful because degeneracy is a source of problems in many mathematical programming algorithms. For example, the class of modified Lagrangian algorithms considered in Fletcher [3] have the property that the Hessian matrix is discontinuous at $\mathbf{x}^*$ if the MPP is degenerate.

2. On the nature of the trajectory in the degenerate case

The results quoted in the previous Section rely heavily on the non-degeneracy assumption. However, the situation when this does not hold has not attracted a great deal of attention. In [2, p. 81] an example is given in which the non-degeneracy condition is not satisfied and in which $\|\Delta_k(\mathbf{x})\| = O(r_k^{\frac{1}{2}})$. In Mifflin [7] it is shown that if the non-degeneracy condition is relaxed in Theorem 1.1, then provided the MPP is strictly convex we have $\|\Delta_k(\mathbf{x})\| \leqslant O(r_k^{\frac{1}{2}})$. Here we improve this result by relaxing the convexity assumptions and show that $\|\Delta_k(\mathbf{x})\| = O(r_k^{\frac{1}{2}})$ whenever $\mathbf{x}^*$ is a regular local solution of a degenerate MPP, and in the Appendix we provide further information on the nature of the convergence of $\mathbf{x}(r_k)$ to $\mathbf{x}^*$ by displaying the asymptotic form of the differential equations governing the trajectory as $r \to 0$ in the degenerate case.

It is instructive to consider the differences between the degenerate and non-degenerate cases in more detail. First recall that if $\mathbf{x}^*$ is a regular local solution and Theorem 1.1 holds then

$$I = I_1, \quad \|\Delta_k(\mathbf{x})\| = O(r_k)$$

and

$$\frac{r_k}{\nabla g_i(\mathbf{x}_k)\Delta_k(\mathbf{x})} \to u_i > 0, \quad \forall\, i \in I_1,$$

where we have written $\mathbf{x}_k$ for $\mathbf{x}(r_k)$. Hence

$$\theta_i = \lim_{r_k \to 0} \cos^{-1}\left(\frac{\nabla g_i(\mathbf{x}_k)\Delta_k(\mathbf{x})}{\|\nabla g_i(\mathbf{x}_k)\|\,\|\Delta_k(\mathbf{x})\|}\right) < \tfrac{1}{2}\pi. \tag{2.1}$$

Now assume that $I_2 \neq \emptyset$. We have $\lim_{k\to\infty}(r_k/g_i(\mathbf{x}_k)) = 0$, $i \in I_2$, so that, as $g_i(\mathbf{x}_k) = \nabla g_i(\mathbf{x}_k)\Delta_k(\mathbf{x}) + o(\|\Delta_k(\mathbf{x})\|)$,

$$r_k = o(\nabla g_i(\mathbf{x}_k)\Delta_k(\mathbf{x}) + o(\|\Delta_k(\mathbf{x})\|)), \quad i \in I_2.$$

Thus, making use of the Cauchy–Schwarz inequality, we obtain

$$r_k = o(\|\Delta_k(\mathbf{x})\|). \tag{2.2}$$

It follows that

$$\frac{|\nabla g_i(\mathbf{x}_k)\Delta_k(\mathbf{x})|}{\|\nabla g_i(\mathbf{x}_k)\|\,\|\Delta_k(\mathbf{x})\|} \to 0, \quad i \in I_1, \tag{2.3}$$

and hence that $\theta_i = \frac{1}{2}\pi$. Thus the sequence $\{\mathbf{x}_k\}$ approaches $\mathbf{x}^*$ along a trajectory that is tangential to the constraint surfaces $g_i(\mathbf{x}) = 0$, $i \in I_1$. It follows that if $\Delta_k(\mathbf{x})$ is decomposed into $\|\Delta_k(\mathbf{x})\|(\mathbf{s}_k + \mathbf{t}_k)$, where $\mathbf{s}_k$ is a linear combination of the $\nabla g_i(\mathbf{x}^*)$, $i \in I_1$, and $\mathbf{t}_k$ lies in the orthogonal complement of this set, then $\|\mathbf{t}_k\| \to 1$ and $\|\mathbf{s}_k\| \to 0$. It is convenient to formalize this result in the following Lemma.

https://doi.org/10.1017/S0334270000001715 Published online by Cambridge University Press

LEMMA 2.1. *Let* $\mathbf{x}^*$ *be a regular local solution,* $I_2 \neq \emptyset$, *and*

$$\Delta_k(\mathbf{x}) = \|\Delta_k(\mathbf{x})\|(\mathbf{t}_k + \mathbf{s}_k),$$

where $\mathbf{s}_k$ *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, then*

(i) *there exists* $\alpha > 0$ *such that* $1 \geqslant \|\mathbf{t}_k\| \geqslant \alpha$ *for* k *large enough,*

(ii) $\|\mathbf{s}_k\| = o(1)$ as $r \to 0$.

We are now in a position to prove the main result of this Section.

THEOREM 2.1. *Let* $\mathbf{x}^*$ *be a regular local solution and* $I_2 \neq \emptyset$, *then we have the asymptotic inequalities*

$$\sqrt{\left(\frac{|I_2|}{M}\right)} r_k^{\frac{1}{2}} + o(r_k^{\frac{1}{2}}) \leqslant \|\Delta_k(\mathbf{x})\| \leqslant \sqrt{\left(\frac{|I_2|}{m\alpha}\right)} r_k^{\frac{1}{2}} + o(r_k^{\frac{1}{2}}) \tag{2.4}$$

with m, M, α *given by equations* (1.5), (1.6) *and Lemma* 2.1.

PROOF. The necessary condition for a minimum of (1.2) at $\mathbf{x}_k$ gives

$$\nabla f(\mathbf{x}_k) - \sum_{i=1}^{m} u_i(r_k) \nabla g_i(\mathbf{x}_k) = 0.$$

Subtracting (1.3) from this system we obtain

$$\nabla_x^2 L(\mathbf{x}^*, \mathbf{u}) \Delta_k(\mathbf{x}) - \sum_{i=1}^{m} \Delta_k(u_i) \nabla g_i(\mathbf{x}^*) = o(\|\Delta_k(\mathbf{x})\|).$$

Multiplying this equation by $\Delta_k(\mathbf{x})^{\mathrm{T}}$ gives

$$\Delta_k(\mathbf{x})^{\mathrm{T}} \nabla_x^2 L(\mathbf{x}^*, \mathbf{u}) \Delta_k(\mathbf{x}) - \sum_{i=1}^{m} \Delta_k(u_i) \Delta_k(g_i) = o(\|\Delta_k(\mathbf{x})\|^2)$$

so that

$$\Delta_k(\mathbf{x})^{\mathrm{T}} \nabla_x^2 L(\mathbf{x}^*, \mathbf{u}) \Delta_k(\mathbf{x}) - \sum_{i \in I_1} \Delta_k(u_i) g_i(\mathbf{x}_k) - \sum_{i \in I_2} u_i(r_k) g_i(\mathbf{x}_k)$$

$$- \sum_{i \in R} \frac{r_k}{g_i(\mathbf{x}_k)} \Delta_k(g_i) = o(\|\Delta_k(\mathbf{x})\|^2).$$

From this equation and Lemma 2.1 it follows that

$$\|\Delta_k(\mathbf{x})\|^2 (\mathbf{t}_k^{\mathrm{T}} \nabla_x^2 L(\mathbf{x}^*, \mathbf{u}) \mathbf{t}_k + o(1)) = r_k(|I_2| + o(1)), \tag{2.5}$$

where we have used that $g_i(\mathbf{x}_k) = O(r_k)$ and $\Delta_k(u_i) \to 0$, $i \in I_1$, and that $g_i(\mathbf{x}_k) > 0$ and $\Delta_k(g_i) = O(\|\Delta_k(\mathbf{x})\|)$ for $i \in R$. Thus (2.4) follows by two easy estimates using the second-order sufficiency conditions, Remark 1.1, and Lemma 2.1.

https://doi.org/10.1017/S0334270000001715 Published online by Cambridge University Press

REMARK 2.1. Lemma 2.1 shows that a degenerate constraint affects considerably the performance of the barrier function algorithms by forcing the solution trajectory against the other active constraints with the resulting reduction in the rate of convergence shown in Theorem 2.1. However, the degenerate constraint is in an important sense redundant as the first-order necessary conditions for a solution of the MPP is unchanged if it is just ignored, and as degeneracy does not affect the second-order conditions. However, it is an indication that the property of membership of the active constraint is extremely sensitive to perturbation of the problem data. Thus numerical algorithms which are not capable of discriminating between the members of the active constrained set could encounter trouble because of this fact. We have illustrated this point for the barrier function method, and believe that it is rightly regarded as a shortcoming. A possible remedy for this particular case has been suggested in [8], and we hope to discuss it further in a companion paper.

3. A measure of degeneracy

In this section our aim is to introduce a measure of non-degeneracy of a MPP. Since for a non-degenerate problem $u_i^* > 0$, $\forall i \in I$, a possible first guess at an appropriate measure is $\max_{i,j} u_i^*/u_j^*$, $i,j \in I$. However, this can be changed arbitrarily by multiplying the constraints by suitably chosen positive numbers. A more suitable choice which is independent of this kind of rescaling of the constraints is

$$\gamma = \max_{i,j} \frac{u_i^* \| \nabla g_i(\mathbf{x}^*) \|}{u_j^* \| \nabla g_j(\mathbf{x}^*) \|}. \tag{3.1}$$

Clearly $\gamma \geqslant 1$ and γ is unbounded as the MPP becomes degenerate. Also the results of the previous section show that in the degenerate case the derivatives $d\mathbf{x}/dr$, $d\mathbf{u}/dr$ are unbounded as $r \to 0$, because of the dependence on $r^{\frac{1}{2}}$. Thus it is natural to ask if there is a link between these two behaviours. The main result of this section confirms such a link and shows that the growth of $d^i\mathbf{x}/dr^i$, $d^i\mathbf{u}/dr^i$ as i increases for fixed r small enough is simply related to γ. To obtain this result, it is convenient first to rescale the problem as follows.

Let

$$\beta = \max_i (u_i^* \| \nabla g_i(\mathbf{x}^*) \|), \tag{3.2}$$

and consider

$$\hat{B}(\mathbf{x}, p) = F(\mathbf{x}) - p \sum_{i=1}^{m} \log G_i(\mathbf{x}), \tag{3.3}$$

where

$$p = r/\beta,$$

$$F(\mathbf{x}) = f(\mathbf{x})/\beta,$$

https://doi.org/10.1017/S0334270000001715 Published online by Cambridge University Press

and

$$G_i(\mathbf{x}) = g_i(\mathbf{x})/\|\nabla g_i(\mathbf{x}^*)\|, \quad i = 1, 2, \ldots, m.$$

Clearly,

$$\nabla_x \hat{B}(\mathbf{x}, p) = \frac{1}{\beta} \nabla_x B(\mathbf{x}, r)$$

so that $\hat{B}$ has the same stationary values as B and the rescaled MPP has the same value of γ. We denote the Lagrange multipliers for the rescaled problems by $\hat{u}_i$, $i = 1, \ldots, m$, and the Lagrangian by $\hat{L}(\mathbf{x}, \hat{\mathbf{u}})$. We have

$$\gamma = \max_{i \in I} \frac{1}{\hat{u}_i}. \tag{3.4}$$

Before constructing the trajectory for the transformed problem it is convenient to prove a preliminary Lemma.

LEMMA 3.1. *Let* $\mathbf{x}^*$ *be a regular local minimum of the (rescaled) MPP and let the matrix* K *be given by*

$$K = \begin{bmatrix} \nabla_x^2 \hat{L}(\mathbf{x}^*, \hat{\mathbf{u}}) & [-\nabla G_i(\mathbf{x}^*)^{\mathrm{T}}, i \in I] & [-\nabla G_i(\mathbf{x}^*)^{\mathrm{T}}, i \in R] \\ [\nabla G_i(\mathbf{x}^*), i \in I] & 0 & 0 \\ 0 & 0 & [\operatorname{diag} G_i(\mathbf{x}^*), i \in R] \end{bmatrix} \tag{3.5}$$

Then K *is non-singular.*

PROOF. Assume K is singular. Then there exists a non-trivial vector $\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ 0 \end{bmatrix}$ such that

$$K \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ 0 \end{bmatrix} = 0.$$

This implies the equations

$$\nabla G_i(\mathbf{x}^*)\,\mathbf{a} = 0, \quad i \in I.$$

Pre-multiplying by $\{\mathbf{a}^{\mathrm{T}}, \mathbf{b}^{\mathrm{T}}, 0\}$ now gives

$$\mathbf{a}^{\mathrm{T}} \nabla_x^2 \hat{L}(\mathbf{x}^*, \hat{\mathbf{u}})\,\mathbf{a} = 0$$

so that the assumption that $\mathbf{x}^*$ is a regular local solution implies that $\mathbf{a} = 0$. But then the linear independence of the set $[\nabla G_i(\mathbf{x}^*)^{\mathrm{T}}, i \in I]$ gives that $\mathbf{b} = 0$ which establishes a contradiction.

We note that this result does not require $I_2 = \emptyset$.

https://doi.org/10.1017/S0334270000001715 Published online by Cambridge University Press

REMARK 3.1. The significance of Lemma 3.1 follows from the fact that in the non-degenerate case we can factor $\hat{J}(0)$ (where $\hat{J}(p)$ is the Jacobian of the rescaled MPP) into the form

$$\hat{J}(0) = \begin{bmatrix} \tilde{I}_n & & \\ & [\operatorname{diag} \hat{u}_i, i \in I] & \\ & & \tilde{I}_R \end{bmatrix} K = QK, \tag{3.6}$$

where $\tilde{I}_n$ and $\tilde{I}_R$ are the $n \times n$ and $|R| \times |R|$ identity matrices, respectively. Consider a family of perturbations of the MPP which increases γ without bound but preserves the regular local solution property of the minimum. It is clear that the approach to degeneracy is captured by the diagonal matrix Q, and that K remains non-singular provided the parameter m in the second-order sufficiency conditions is bounded away from zero (as it must be if the minimum remains a regular local solution). Thus we expect an estimate of $\|\hat{J}(0)^{-1}\|$ of the form

$$\|\hat{J}(0)^{-1}\| = \kappa \frac{\gamma}{m}, \tag{3.7}$$

where κ is an (order 1) constant.

REMARK 3.2. A convenient way to construct such a family of problems is to start with a degenerate MPP having a regular local solution at $\mathbf{x}^*$ and then to modify the objective function by adding a function $w(\mathbf{x}, \boldsymbol{\lambda})$ given by

$$w(\mathbf{x}, \boldsymbol{\lambda}) = \sum_{i \in I_2} \lambda_i g_i(\mathbf{x}), \quad \lambda_i \geq 0.$$

The effect of $w(\mathbf{x}, \boldsymbol{\lambda})$ is such that the multiplier u_i^* is unchanged for $i \notin I_2$, and $u_i^* = \lambda_i$ for $i \in I_2$. Clearly the λ_i can be chosen such that γ takes any value ≥ 1 while the contributions of $\nabla_x^2 w(\mathbf{x}, \boldsymbol{\lambda})$ and $\sum_{i \in I_2} \lambda_i \nabla^2 g_i(\mathbf{x})$ are such that $\nabla_x^2 L$ remains unchanged. Also the second-order conditions continue to hold with the same value of m as the set of allowable vectors $\mathbf{t}$ in (1.5) is further constrained if any $\lambda_i > 0$. From this family we can choose a sequence of problems having the property that γ becomes unbounded while the regular local solution properties holds uniformly.

THEOREM 3.1. *Consider a family of MPP's indexed by a parameter α having the following properties.*

(i) *For each α, $\mathbf{x}^*$ is a regular local solution,*

(ii) *there exists a constant $m > 0$ such that* (1.5) *holds for this m and each α, and*

(iii) $\lim_{\alpha \to \infty} \gamma(\alpha) = +\infty$,

https://doi.org/10.1017/S0334270000001715 Published online by Cambridge University Press

then for fixed p small enough we have the asymptotic estimates

$$\| D^{q+1}\mathbf{x}_\alpha\| = O(\gamma(\alpha)^{2q+1}), \quad \| D^{q+1}\hat{\mathbf{u}}_\alpha\| = O(\gamma(\alpha)^{2q+1}) \tag{3.8}$$

where $D \equiv d/dp$, *and* $\mathbf{x}_\alpha(p)$, $\hat{\mathbf{u}}_\alpha(p)$ *are defined by minimizing* (3.3) *for the* α*th member of the family of MPP's.*

PROOF. It is convenient to drop the subscript α and to write $\mathbf{z} = \begin{bmatrix}\mathbf{x}\\ \hat{\mathbf{u}}\end{bmatrix}$. From equation (1.8) it follows that

$$\hat{J}D\mathbf{z} = QKD\mathbf{z} = \begin{bmatrix}0\\ \mathbf{e}\end{bmatrix} \tag{3.9}$$

so that, as K is non-singular by assumption,

$$\| D\mathbf{z}\| = O(\| Q^{-1}\|) = O(\gamma),$$

and this establishes the result in the special case $q = 0$. In fact the estimate holds also for $DA(\mathbf{z})$ where A is any smooth enough function of $\mathbf{z}$ for by the chain rule

$$|DA(\mathbf{z})| \leqslant \left\| \frac{\partial A(\mathbf{z})}{\partial \mathbf{z}} \right\| \| D\mathbf{z}\|,$$

where

$$\frac{\partial A(\mathbf{z})}{\partial \mathbf{z}} = \left[\frac{\partial A}{\partial z_1}, \quad \frac{\partial A}{\partial z_2}, \quad \dots\right].$$

The proof now follows by induction. We assume that

$$|D^l A(\mathbf{z})| = O(\gamma^{2l-1}), \quad l \leqslant q, \tag{3.10}$$

and show that this implies that $\| D^{q+1}\mathbf{z}\| = O(\gamma^{2q+1})$. Differentiating (3.9) gives

$$\hat{J}D^{q+1}\mathbf{z} = -\sum_{l=1}^{q}\binom{q}{l} D^l\hat{J}D^{q+1-l}\mathbf{z}. \tag{3.11}$$

To estimate the order of each term on the right-hand side we apply the induction hypothesis (3.10). This gives an order of

$$(2l-1)+(2q+2-2l-1) = 2q$$

for each l. The estimate for $D^{q+1}\mathbf{z}$ follows as an extra order in γ must be counted for the inversion of $\hat{J}$ in (3.11). To complete the induction it is necessary to show that (3.10) holds with $q = q+1$ as a consequence of the estimate established for

https://doi.org/10.1017/S0334270000001715 Published online by Cambridge University Press

$D^{q+1}\mathbf{z}$. We have

$$D^{q+1}A(\mathbf{z}) = D^q\left[\frac{\partial A(\mathbf{z})}{\partial \mathbf{z}} \quad D\mathbf{z}\right]$$

$$= \sum_{i=0}^{q}\binom{q}{i} D^i \frac{\partial A(\mathbf{z})}{\partial \mathbf{z}} D^{q+1-i}\mathbf{z}.$$

For $i = 0$ the corresponding term is the scalar product of $\partial A(\mathbf{z})/\partial \mathbf{z}$ and $D^{q+1}\mathbf{z}$ so that it necessarily has order $2q+1$ in γ. For $i>0$ the order estimate is a consequence of (3.10) which gives

$$(2i-1)+2(q+1-i)-1 = 2q.$$

This is of smaller order than the term for $i = 0$ so that

$$|D^{q+1}A(\mathbf{z})| = O(\gamma^{2q+1}).$$

Thus the induction hypothesis is verified and the Theorem follows.

REMARK 3.3. An appropriate formalism for developing expressions of the form (3.11) into their component parts has been developed by Butcher in a number of papers (for example [1]) for his work on Runge–Kutta methods. One deduction from this work is that the numerical coefficients hidden in the order estimates (3.8) are likely to grow like $q!$ while (3.7), (3.9) suggests that a term $(1/m)^{q+1}$ is also likely to be present. It is convenient to assume a dependence of this form for the numerical coefficient. However, it is not critical to the development of our arguments.

4. The numerical performance of extrapolation

If we combine the results of the previous Section with the formula (1.9) for the error in a linear extrapolation procedure, then it follows that the error in an extrapolation based on the points $p = p_{k-s+1}, \ldots, p_k$ is proportional to

$$e_{k,s} = \left(\prod_{i=k-s+1}^{k} p_i\right)\gamma^{2s-1}, \tag{4.1}$$

where we have omitted any explicit dependence of a coefficient on s as a result of Remark 3.3, and any dependence on m as a "second best" compromise because its estimation is relatively more difficult. The numerical situation we consider is one in which we have successive minima $\mathbf{x}(p_1), \ldots, \mathbf{x}(p_k)$, and where we have carried out an extrapolation on the points $p = p_{k-s+1}, \ldots, p_k$. We ask is it worth-while to include data from the point $p = p_{k-s}$ in the extrapolation, and as a basis for this

https://doi.org/10.1017/S0334270000001715 Published online by Cambridge University Press

decision we have tested the heuristic

$$\mu_{k,s} = \frac{e_{k,s+1}}{e_{k,s}} = p_{k-s}\gamma^2 < 0.1.$$

Hopefully, if the strategy is successful and if the terms in $e_{k,s}$ actually are the most significant, then the new extrapolation would produce at least one more decimal place of accuracy. The performance in practice has been surprisingly successful, giving accurate predictions for a range of test problems including the "easy" Rosen–Suzuki problem and a range of the Colville test problems including the "Shell dual" (Colville II) which is generally regarded as quite difficult and which proves to have a large value of γ.

The basic numerical data for the tabulated results is summarized in Table 4.1. The sequence of minimizations is designed to reproduce the results of Fletcher and McCann [4]. In Tables 4.2 and 4.3 we give $\max_{i \in I}|g_i(\mathbf{x})|$ for $\mathbf{x}$ resulting from the extrapolation procedure. The minimizations have been carried out using a severe error tolerance to ensure that extraneous errors should not unduly perturb the results. It would seem that we have achieved results accurate to at least 12 decimal places in both cases.

TABLE 4.1

Basic data for the numerical results

Test problem	γ	β	p_i
Rosen–Suzuki	1.6	8	1.2×10^{-i}
Colville II	83	57	0.16×10^{-i}

For the Rosen–Suzuki problem the heuristic is satisfied provided $k-s \geqslant 2$ suggesting that repeated extrapolation is justified on the set of points $p = p_2, \ldots, p_k$, and this is borne out very well by Table 4.2. For Colville II the heuristic is satisfied provided the extrapolation is based on the points $p = p_4, \ldots, p_k$, and the results in Table 4.3 again support this strategy.

REMARK 4.1. Tables 4.2 and 4.3 both show that there is a tendency for the extrapolation to improve accuracy until a threshold is reached and, from this point on, this threshold is essentially preserved. This confirms that extrapolation with this choice of parameters is a very stable process. Going down the table, the threshold produced is a function of the accuracy of the successive minimizations and depends quite strongly on the accuracy of the final maximization.

https://doi.org/10.1017/S0334270000001715 Published online by Cambridge University Press

TABLE 4.2

Numerical results for Rosen–Suzuki problem

k \ s	0	1	2	3	4	5	6	7	8	9
1	1.1									
2	1.1×10^{-1}	6.0×10^{-3}								
3	1.0×10^{-2}	1.4×10^{-3}	1.3×10^{-3}							
4	1.0×10^{-3}	1.9×10^{-5}	5.2×10^{-6}	3.8×10^{-6}						
5	1.0×10^{-4}	1.9×10^{-7}	5.9×10^{-9}	7.2×10^{-10}	3.4×10^{-10}					
6	1.0×10^{-5}	2.0×10^{-9}	7.2×10^{-12}	1.3×10^{-12}	1.2×10^{-12}	1.2×10^{-12}				
7	1.0×10^{-6}	2.1×10^{-11}	9.8×10^{-13}	9.8×10^{-13}	9.8×10^{-13}	9.8×10^{-13}	9.8×10^{-13}			
8	1.0×10^{-7}	3.0×10^{-13}	5.2×10^{-13}	5.2×10^{-13}	5.2×10^{-13}	5.2×10^{-13}	5.2×10^{-13}	5.2×10^{-13}		
9	1.0×10^{-8}	5.2×10^{-13}	5.2×10^{-13}	5.2×10^{-13}	5.2×10^{-13}	5.2×10^{-13}	5.2×10^{-13}	5.2×10^{-13}	5.2×10^{-13}	
10	1.0×10^{-9}	5.8×10^{-14}	6.2×10^{-14}	6.2×10^{-14}	6.2×10^{-14}	6.2×10^{-14}	6.2×10^{-14}	6.2×10^{-14}	6.2×10^{-14}	6.2×10^{-14}

https://doi.org/10.1017/S0334270000001715 Published online by Cambridge University Press

TABLE 4.3

Numerical results for Colville II

k \ s	0	1	2	3	4	5	6	7	8	9
1	4.2									
2	3.7×10^{-1}	1.6×10^{-1}								
3	4.3×10^{-2}	8.5×10^{-3}	7.0×10^{-3}							
4	4.4×10^{-3}	2.7×10^{-4}	1.9×10^{-4}	1.8×10^{-4}						
5	4.5×10^{-4}	4.4×10^{-6}	1.8×10^{-6}	1.6×10^{-6}	1.6×10^{-6}					
6	4.5×10^{-5}	4.8×10^{-8}	4.2×10^{-9}	2.4×10^{-9}	2.3×10^{-9}	2.3×10^{-9}				
7	4.5×10^{-6}	4.8×10^{-10}	1.0×10^{-11}	9.9×10^{-12}	9.8×10^{-12}	9.8×10^{-12}	9.8×10^{-12}			
8	4.5×10^{-7}	1.3×10^{-11}	8.7×10^{-12}	8.7×10^{-12}	8.7×10^{-12}	8.7×10^{-12}	8.7×10^{-12}	8.7×10^{-12}		
9	4.5×10^{-8}	7.4×10^{-12}	7.4×10^{-12}	7.4×10^{-12}	7.4×10^{-12}	7.4×10^{-12}	7.4×10^{-12}	7.4×10^{-12}	7.4×10^{-12}	
10	4.5×10^{-9}	4.1×10^{-12}	4.1×10^{-12}	4.1×10^{-12}	4.1×10^{-12}	4.1×10^{-12}	4.1×10^{-12}	4.1×10^{-12}	4.1×10^{-12}	4.1×10^{-12}

REMARK 4.2. The success of the heuristic for the Colville II problem suggests that near degeneracy might well be the source of the numerical problems reported, and this raises the possibility that the difficulties could be the result of algorithmic shortcomings rather than a consequence of any particular difficulty inherent in the problem.

Acknowledgement

We are indebted to a referee for helpful comments which have led to improvements in presentation.

Appendix

Let $\mathbf{x}^*$ be a regular local solution for the MPP, and $I_2 \neq \emptyset$. For r_k small enough it follows from the second-order sufficiency condition that $\nabla_x^2 B(\mathbf{x}_k, r_k)$ is positive definite. Therefore there exists $\delta > 0$ such that, by the implicit function theorem, the system of equations $\nabla_x B(\mathbf{x}, r) = 0$ defines a smooth trajectory $\mathbf{x} = \mathbf{x}(r)$ on the open interval $0 < r < \delta$, and, by assumption, $\lim_{r \to 0} \mathbf{x}(r) = \mathbf{x}^*$.

LEMMA A.1. *Let the possible partitions of I_2 into two disjoint sets be ordered with respect to an index i such that*

$$P_i \cup Q_i = I_2, \quad i = 1, 2, \ldots, \sigma. \tag{A.1}$$

Then

$$\det(J(r)) = \sum_{i=1}^{\delta} A_i(r) \prod_{j \in P_i} u_j(r) \prod_{j \in Q_i} g_j(\mathbf{x}(r)), \tag{A.2}$$

where

$$A_i(r) = \theta_i \det \left(\begin{bmatrix} \nabla_x^2 L(\mathbf{x}, \mathbf{u}) & [-\nabla g_j(\mathbf{x})^{\mathrm{T}}, j \in I_1] & [-\nabla g_j(\mathbf{x})^{\mathrm{T}}, j \in R] & [-\nabla g_j(\mathbf{x})^{\mathrm{T}}, j \in P_i] \\ [u_j \nabla g_j(\mathbf{x}), j \in I_1] & [g_j(\mathbf{x}), j \in I_1] & 0 & 0 \\ [u_j \nabla g_j(\mathbf{x}), j \in R] & 0 & [g_j(\mathbf{x}), j \in R] & 0 \\ [\nabla g_j(\mathbf{x}), j \in P_i] & 0 & 0 & 0 \end{bmatrix} \right), \tag{A.3}$$

$\theta_i = \pm 1$ *and* $A_i(0)$ *is non-zero*, $i = 1, 2, \ldots$.

https://doi.org/10.1017/S0334270000001715 Published online by Cambridge University Press

PROOF. Let $k \in I_2$. Then, by suitably interchanging rows and columns, we can write

$$\theta \det J(r) = \det \left(\begin{array}{cc|c} & M & -\nabla g_k^{\mathrm{T}} \\ & & 0 \\ \hline u_k & \nabla g_k \quad 0 & g_k \end{array} \right)$$

$$= u_k \det \left(\begin{array}{cc|c} & M & -\nabla g_k^{\mathrm{T}} \\ & & 0 \\ \hline \nabla g_k & 0 & 0 \end{array} \right) + g_k \det M,$$

where $\theta = \pm 1$. The result (A.2) now follows by repeated applications of this device. That $A_i(0)$ is non-zero follows by the argument used to establish Lemma 3.1.

COROLLARY. *$J(r)$ is non-singular for $r > 0$, small enough.*

The trajectory $\mathbf{x} = \mathbf{x}(r)$ must satisfy the system of differential equations

$$J(r) \begin{bmatrix} d\mathbf{x}/dr \\ d\mathbf{u}/dr \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{e} \end{bmatrix}. \tag{A.4}$$

It follows from (A.2) that this system becomes singular as $r \to 0$ when $I_2 \neq \emptyset$, and our main concern here is to show how this singular behaviour may be isolated. Equation (A.4) can be written in explicit form

$$\begin{bmatrix} d\mathbf{x}/dr \\ d\mathbf{u}/dr \end{bmatrix} = \frac{\text{adjoint}(J)}{\det(J)} \begin{bmatrix} 0 \\ \mathbf{e} \end{bmatrix} = \frac{1}{\det(J)} \begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ \sum_{\alpha=1}^{m} (-1)^{\alpha+\beta} |J_{\alpha\beta}| \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}, \tag{A.5}$$

where $|J_{\alpha\beta}|$ is the minor obtained by eliminating the row associated with constraint α and the column associated with constraint β. By considering the form of expansion leading to (A.2) it will be seen that the dominant terms for small r will be associated with $\alpha \in I_2$, and we concentrate on the equations corresponding to $\beta \in I_2$. To evaluate $|J_{\alpha\beta}|$ we note that there are two cases:

(i) $\alpha = \beta$. Let $I_\beta = I_2 - \{\beta\}$, and order all possible partitions of I_β into disjoint sets P_i^β, Q_i^β, $P_i^\beta \cup Q_i^\beta = I_\beta$, with $1 \leqslant i \leqslant \sigma_1$. Note that for each i there exists $j(i, \beta)$ such that

$$P_i^\beta = P_{j(i,\beta)}, \tag{A.6}$$

https://doi.org/10.1017/S0334270000001715 Published online by Cambridge University Press

and this identification permits us to express $|J_{\beta\beta}|$ in the form

$$|J_{\beta\beta}| = \sum_{i=1}^{\sigma_1} \prod_{k\in P_i^\beta} u_k \prod_{k\in Q_i^\beta} g_k A_{j(i,\beta)}. \tag{A.7}$$

(ii) $\alpha\neq\beta$. Let $I_{\alpha\beta} = I_2 - \{\alpha,\beta\}$ and order all possible partitions of $I_{\alpha\beta}$ into disjoint sets $P_i^{\alpha\beta}$, $Q_i^{\alpha\beta}$, $P_i^{\alpha\beta}\cup Q_i^{\alpha\beta} = I_{\alpha\beta}$, with $1<i<\sigma_2$. In this case the expansion procedure gives

$$|J_{\alpha\beta}| = u_\beta \sum_{i=1}^{\sigma_2} \prod_{k\in P_i^{\alpha\beta}} u_k \prod_{k\in Q_i^{\alpha\beta}} g_k A_i^{\alpha\beta}, \tag{A.8}$$

where $A_i^{\alpha\beta}$ has the form

$$A_i^{\alpha\beta} =$$

$$\theta_i \det \begin{pmatrix} \nabla_x^2 L & [-\nabla g_j^{\mathrm{T}}, j\in I_1] & [-\nabla g_j^{\mathrm{T}}, j\in R] & -\nabla g_\alpha^{\mathrm{T}} & [-\nabla g_j^{\mathrm{T}}, j\in P_i^{\alpha\beta}] \\ [u_j \nabla g_j, j\in I_1] & [g_j, j\in I_1] & 0 & 0 & 0 \\ [u_j \nabla g_j, j\in R] & 0 & [g_j, j\in R] & 0 & 0 \\ \nabla g_\beta & 0 & 0 & 0 & 0 \\ [\nabla g_j, j\in P_i^{\alpha\beta}] & 0 & 0 & 0 & 0 \end{pmatrix} \tag{A.9}$$

and $\theta_i = \pm 1$. In this case it is not true that $A_i^{\alpha\beta}(0)\neq 0$ necessarily.

THEOREM A.1. *If* $u_\alpha(r) = c_\alpha r^q + o(r^q)$, $\alpha\in I_2$, *as* $r\to 0$, *then necessarily* $q = \frac{1}{2}$.

PROOF. Note that $u_\alpha(r) = r/g_\alpha(\mathbf{x}(r))\to 0$, $r\to 0$ implies that $q<1$. Substituting for u_α, $\alpha\in I_2$ in the equation for du_β/dr and using (A.2, A.5, A.7, A.8) gives

$$\frac{du_\beta}{dr} = \frac{c}{r^q} + \textit{smaller terms in } r, \quad \beta\in I_2,$$

where c is a non-zero constant, and the result follows from this.

REMARK A.1. This result complements Theorem 2.1 in which it was shown that $\|\Delta_k(\mathbf{x})\| = O(r_k^{\frac{1}{2}})$.

References

[1] J. C. Butcher, "Coefficients for the study of Runge–Kutta integration processes", *J. Aust. Math. Soc.* 8 (1963), 185–201.

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

https://doi.org/10.1017/S0334270000001715 Published online by Cambridge University Press

[3] R. Fletcher, "An ideal penalty function for constrained optimization", *J. Inst. Maths. Applics.* 15 (1975), 319–342.
[4] R. Fletcher and A. P. McCann, "Acceleration techniques for nonlinear programming", in *Numerical methods for nonlinear optimization* (ed. F. A. Lootsma) (Academic Press, London, 1972), pp. 203–214.
[5] P. J. Laurent, "Convergence du procédé d'extrapolation de Richardson", *Troisième Congrès de l'Afcalti* (Toulouse), pp. 81–98.
[6] F. A. Lootsma, "Extrapolation in logarithmic programming", *Philips Res. Repts* 23 (1968), 108–116.
[7] R. Mifflin, "Convergence bounds for nonlinear programming algorithms", *Math. Programming* 8 (1975), 251–271.
[8] M. R. Osborne, "Topics in optimizations", *Comp. Sci. Dept, Stanford Univ.* (STAN–CS–72–279, 1972).

Computer Centre
Australian National University
P.O. Box 4,
Canberra, A.C.T. 2600