INT. J. SYSTEMS SCI., 1976, VOL. 7, NO. 11, 1239–1248

Non-linear programming using least *p*th optimization with extrapolation†

J. W. BANDLER‡ and W. Y. CHU§

We present a general approach for solving minimax and non-linear programming problems through a sequence of least *p*th approximations with extrapolation. Several numerical examples illustrate the effectiveness of the present approach. A comparison with the well-known SUMT method of Fiacco and McCormick is made under the same conditions and employing Fletcher's quasi-Newton programme.

1. Introduction

It is well known that least *p*th approximation with a very large value of p can, in principle, be used to achieve a near minimax solution (Bandler 1969, Bandler and Charalambous 1972, 1973). For numerical efficiency, the process may be accomplished by using a sequence of least *p*th approximations with increasing values of p. By this approach, a sequence of least *p*th minima will be obtained. Under appropriate assumptions we may expect the sequence of least *p*th minima to form a unique trajectory of local minima converging to the minimax optimum, and the extrapolation technique used by Fiacco and McCormick (1968) and Lootsma (1968) may be applied to accelerate convergence. Several numerical examples are used to illustrate the effectiveness of the extrapolation technique applied to least *p*th approximations. Theoretical validation of the new approach is also given.

Using the Bandler–Charalambous (1974) minimax formulation we can readily transform a non-linear programming problem into a minimax problem to be solved by the present approach.

2. Basic formulae

A brief review of the formulae used in solving the test examples will be presented.

2.1. *Generalized least pth objective*

The generalized least *p*th objective function (Bandler and Charalambous 1972) to be minimized with respect to $\boldsymbol{\phi}$ is

$$U(\boldsymbol{\phi},p)=\begin{cases} M(\boldsymbol{\phi})\left(\sum\limits_{i\in K}\left(\dfrac{e_i(\boldsymbol{\phi})}{M(\boldsymbol{\phi})}\right)^q\right)^{1/q} & \text{for } M(\boldsymbol{\phi})\neq 0 \\ 0 & \text{for } M(\boldsymbol{\phi})=0 \end{cases} \tag{1}$$

where $e_i(\boldsymbol{\phi})$ is a set of $m+1$ real error functions, $\boldsymbol{\phi} \triangleq [\phi_1 \phi_2 \cdots \phi_n]^{\mathrm{T}}$

Received 9 December 1975.

†This paper is based on work presented at the *12th Annual Allerton Conference on Circuit and System Theory*, Monticello, Illinois, 1974. This work was supported by the National Research Council of Canada through a scholarship to W.Y. Chu and under Grant A7239, and by the Defence Research Board of Canada under Grant 9931–39.

‡ Group on Simulation, Optimization and Control and Department of Electrical Engineering, McMaster University, Hamilton, Canada L8S 4L7.

§ Bell–Northern Research, Ottawa, Ontario, Canada.

$$q \triangleq p \operatorname{sgn} M(\boldsymbol{\phi}),\ 1 < p < \infty \tag{2}$$

$$M(\boldsymbol{\phi}) \triangleq \max_{i \in I} e_i(\boldsymbol{\phi}) \tag{3}$$

and

$$K = \begin{cases} I \triangleq \{1, 2, \ldots, m+1\} & \text{if } M(\boldsymbol{\phi}) < 0 \\ J \triangleq \{i \mid e_i(\boldsymbol{\phi}) > 0,\ i \in I\} & \text{if } M(\boldsymbol{\phi}) > 0 \end{cases} \tag{4}$$

The gradient vector of the objective function is given by

$$\nabla U(\boldsymbol{\phi}, p) = \left(\sum_{i \in K} \left(\frac{e_i(\boldsymbol{\phi})}{M(\boldsymbol{\phi})} \right)^q \right)^{1/q-1} \sum_{i \in K} \left(\frac{e_i(\boldsymbol{\phi})}{M(\boldsymbol{\phi})} \right)^{q-1} \nabla e_i(\boldsymbol{\phi}) \ \text{for } M(\boldsymbol{\phi}) \neq 0 \tag{5}$$

From (1) and (5) we note that if the $e_i(\boldsymbol{\phi})$ are continuous with continuous first partial derivatives, then, under the stated conditions, the objective function is continuous everywhere with continuous first partial derivatives (except possibly when $M(\boldsymbol{\phi}) = 0$ and two or more maxima are equal).

2.2. *Minimax approach to non-linear programming*

The non-linear programming problem of minimizing $f(\boldsymbol{\phi})$ subject to

$$g_i(\boldsymbol{\phi}) \geqslant 0,\ i = 1, 2, \ldots, m \tag{6}$$

can be transformed into the following unconstrained objective (Bandler and Charalambous 1974):

$$V(\boldsymbol{\phi}, \alpha) = \max_{1 \leqslant i \leqslant m} [f(\boldsymbol{\phi}), f(\boldsymbol{\phi}) - \alpha g_i(\boldsymbol{\phi})] \tag{7}$$

where α is positive, satisfying

$$\frac{1}{\alpha} \sum_{i=1}^{m} \breve{u}_i < 1 \tag{8}$$

where the $\breve{u}_i$s are the Kuhn–Tucker multipliers at the optimum. The minimization of $V(\boldsymbol{\phi}, \alpha)$ with respect to $\boldsymbol{\phi}$ is a minimax problem and may be solved, for example, by minimizing the generalized least pth objective with

$$e_i(\boldsymbol{\phi}) \triangleq f(\boldsymbol{\phi}) - \alpha g_i(\boldsymbol{\phi}),\ i = 1, 2, \ldots, m \tag{9}$$

$$e_{m+1}(\boldsymbol{\phi}) \triangleq f(\boldsymbol{\phi}) \tag{10}$$

using a very large value of p or a sequence of p values with extrapolation. We note that a feasible starting point is not required.

2.3. *Extrapolation polynomials* (*Fiacco and McCormick* 1968)

Suppose the generalized least pth objective function $U(\boldsymbol{\phi}, p)$ is uniquely minimized for $1 < p_1 < \ldots < p_k < \infty$ at $\boldsymbol{\phi}(1/p_1), \ldots, \boldsymbol{\phi}(1/p_k)$. Let $p' \triangleq 1/p$. A polynomial in p' that yields $\boldsymbol{\phi}(p_1'), \ldots, \boldsymbol{\phi}(p_k')$ is given by

$$\boldsymbol{\phi}(p_i') = \sum_{j=0}^{k=1} \mathbf{a}_j (p_i')^j,\ i = 1, \ldots, k \tag{11}$$

where the $\mathbf{a}_j$ are n-component vectors. The determinant of the matrix of coefficients is the Vandermonde determinant and is non-zero if $p_i' \neq p_j'$ for $i \neq j$, in which case we have a unique solution for the $\mathbf{a}_j$. Then $\sum_{j=0}^{k-1} \mathbf{a}_j(p')^j$ is an approximation of $\boldsymbol{\phi}(p')$ on $[0, p_1']$, and $\boldsymbol{\phi}(0) = \check{\boldsymbol{\phi}}$ (the minimax solution) is approximated by $\mathbf{a}_0$.

Now, the exact Taylor series expansion of $\boldsymbol{\phi}(p_i')$ in p_i' about $\boldsymbol{\phi}(0)$ is

$$\boldsymbol{\phi}(p_i') = \sum_{j=0}^{k-1} (p_i')^j \frac{D^j \boldsymbol{\phi}(0)}{j!} + \boldsymbol{\epsilon}^i, i = 1, \ldots, k \tag{12}$$

where

$$D\boldsymbol{\phi}(p') \triangleq \left[\frac{d\phi_1(p')}{dp'} \quad \ldots \quad \frac{d\phi_n(p')}{dp'} \right]^{\mathrm{T}} \tag{13}$$

and $\boldsymbol{\epsilon}^i$ is an error term. It can be shown that the difference between $\mathbf{a}_0$ and $\boldsymbol{\phi}(0)$ is of the order of $(p_1')^k$. Thus, as $p_1' \to 0$, $\mathbf{a}_0 \to \boldsymbol{\phi}(0)$. In addition, the estimates using k minima are better than those using $k-1$ minima. With $p_{i+1}' = p_i'/c$ $(c > 1)$, the particular structure of these equations renders the use of an extrapolation procedure according to the Richardson–Romberg principle (Joyce 1971) to estimate $\mathbf{a}_0$.

If $\boldsymbol{\phi}^{ji}$, $i = 1, \ldots, k$, $j = 1, \ldots, i-1$ signifies the jth-order estimate of $\boldsymbol{\phi}(0)$ after i minima have been obtained, with p_1' being the initial value of p', then we have

$$\boldsymbol{\phi}_0^{\,i} = \boldsymbol{\phi}\left(\frac{p_1'}{c^{i-1}}\right), \quad i = 1, \ldots, k$$

and

$$\boldsymbol{\phi}_j^{\,i} = \frac{c^j \boldsymbol{\phi}_{j-1}^{\,i} - \boldsymbol{\phi}_{j-1}^{\,i-1}}{c^j - 1}, \quad \begin{array}{l} i = 2, \ldots, k \\ j = 1, \ldots, i-1 \end{array} \tag{14}$$

The ' best ' estimate of $\boldsymbol{\phi}(0)$, namely $\mathbf{a}_0$, is given by

$$\boldsymbol{\phi}(0) \cong \boldsymbol{\phi}_{k-1}^{\,k} = \mathbf{a}_0 \tag{15}$$

The extrapolation formula (14) can also be used to estimate the next minimum of the objective function $U(\boldsymbol{\phi}, p)$, i.e. the $(k+1)$th minimum. Setting $i = k+1$ in (14) and solving for $\boldsymbol{\phi}_{j-1}^{\,k+1}$, we have

$$\boldsymbol{\phi}_{j-1}^{\,k+1} = \frac{(c^j - 1)\boldsymbol{\phi}_j^{\,k+1} + \boldsymbol{\phi}_{j-1}^{\,k}}{c^j} \tag{16}$$

Noting that $\mathbf{a}_0 = \boldsymbol{\phi}_{k-1}^{\,k} = \boldsymbol{\phi}_{k-1}^{\,k+1}$ from (15) and using the values previously obtained from (14), we can evaluate (16) for $j = k-1, k-2, \ldots, 1$. The last computation will give the required estimate $\boldsymbol{\phi}_0^{\,k+1}$. This estimate can be used as the starting-point for the $(k+1)$th minimization of $U(\boldsymbol{\phi}, p)$. As more minima are achieved, the estimate eventually improves. This accelerates the entire process by substantially reducing the effort required to minimize the successive U functions.

3. Theoretical justification

We require an isolated trajectory of least pth minima which is a continuously differentiable function in p' for $1 > p' \geqslant 0$ and therefore can be expanded as a Taylor series about $p' = 0$. To justify this we assume

(A 1) The error functions $e_i(\boldsymbol{\phi})$ for $i \in I$ are convex and have continuous $(k+1)$th order, $k \geqslant 1$, partial derivatives with respect to $\boldsymbol{\phi}$.

(A 2) The Hessian matrix of the objective function U is non-singular in the region $\{\boldsymbol{\phi} | M(\boldsymbol{\phi})/M(\check{\boldsymbol{\phi}}) > 0\}$ for every $1 > p' > 0$.

(A 3) Assumptions (developed later) to ensure differentiability of the trajectory at $p = \infty$.

At the minimizing point $\boldsymbol{\phi}(p')$ we have

$$\nabla U(\boldsymbol{\phi}(p'), p) = \left(\sum_{i \in K} \left(\frac{e_i(\boldsymbol{\phi}(p'))}{M(\boldsymbol{\phi}(p'))} \right)^q \right)^{1/q - 1} \sum_{i \in K} \left(\frac{e_i(\boldsymbol{\phi}(p'))}{M(\boldsymbol{\phi}(p'))} \right)^{q-1} \nabla e_i(\boldsymbol{\phi}(p')) = \mathbf{0} \tag{17}$$

Since by assumption the Hessian matrix of U is non-singular, the implicit function theorem assures us that $\boldsymbol{\phi}(p')$ is a continuously differentiable vector function of p' for $1 > p' > 0$. In other words, we have an isolated trajectory of unconstrained local minima of U.

It is possible to be explicit about the derivatives of $\boldsymbol{\phi}(p')$ with respect to p' for $1 > p' > 0$. For convenience, let

$$e_{ip'} \triangleq e_i(\boldsymbol{\phi}(p')) \tag{18}$$

$$M_{p'} \triangleq M(\boldsymbol{\phi}(p')) \tag{19}$$

Since (17) is an identity in $1/q$ (or rather p'), we can differentiate with respect to p', obtaining

$$\left(\sum_{i \in K} \left(\frac{e_{ip'}}{M_{p'}} \right)^q \right)^{1/q-1} \left\{ \sum_{i \in K} \left(\frac{e_{ip'}}{M_{p'}} \right)^{q-1} \nabla (\nabla e_{ip'})^{\mathrm{T}} D\boldsymbol{\phi}(p') + (q-1) \left(\frac{e_{ip'}}{M_{p'}} \right)^{q-2} \left(\frac{\nabla e_{ip'}}{M_{p'}} \right) (\nabla e_{ip'})^{\mathrm{T}} D\boldsymbol{\phi}(p') - (\operatorname{sgn} M_{p'}) q^2 \left(\frac{e_{ip'}}{M_{p'}} \right)^{q-1} \ln \left(\frac{e_{ip'}}{M_{p'}} \right) \nabla e_{ip'} \right\} = \mathbf{0} \tag{20}$$

Now

$$\nabla (\nabla U(\boldsymbol{\phi}(p'), p))^{\mathrm{T}} = \left(\sum_{i \in K} \left(\frac{e_{ip'}}{M_{p'}} \right)^q \right)^{1/q-1} \sum_{i \in K} \left\{ \left(\frac{e_{ip'}}{M_{p'}} \right)^{q-1} \nabla (\nabla e_{ip'})^{\mathrm{T}} + (q-1) \left(\frac{e_{ip'}}{M_{p'}} \right)^{q-2} \left(\frac{\nabla e_{ip'}}{M_{p'}} \right) (\nabla e_{ip'})^{\mathrm{T}} \right\} \tag{21}$$

Equation (20) can hence be written as

$$\nabla (\nabla U(\boldsymbol{\phi}(p'), p))^{\mathrm{T}} D\boldsymbol{\phi}(p') - (\operatorname{sgn} M_{p'}) q^2 \left(\sum_{i \in K} \left(\frac{e_{ip'}}{M_{p'}} \right)^q \right)^{1/q-1} \times \sum_{i \in K} \left\{ \left(\frac{e_{ip'}}{M_{p'}} \right)^{q-1} \ln \left(\frac{e_{ip'}}{M_{p'}} \right) \nabla e_{ip'} \right\} = \mathbf{0} \tag{22}$$

where $D\boldsymbol{\phi}(p')$ was defined in (13). Hence,

$$D\boldsymbol{\phi}(p') = (\operatorname{sgn} M_{p'})\left\{ \nabla(\nabla U(\boldsymbol{\phi}(p'), p))^{\mathrm{T}} \right\}^{-1} q^2 \left(\sum_{i \in K} \left(\frac{e_{ip'}}{M_{p'}} \right)^q \right)^{1/q-1}$$

$$\times \sum_{i \in K} \left\{ \left(\frac{e_{ip'}}{M_{p'}} \right)^{q-1} \ln \left(\frac{e_{ip'}}{M_{p'}} \right) \nabla e_{ip'} \right\} \tag{23}$$

If we differentiate (22) with respect to p', we shall find that the existence of the same inverse is required for $D^2\boldsymbol{\phi}(p')$ to exist as required for $D\boldsymbol{\phi}(p')$. In addition, $D^2\boldsymbol{\phi}(p')$ requires the existence of the third partial derivatives of $e_i(\boldsymbol{\phi})$ with respect to $\boldsymbol{\phi}$. By continuing in this manner it should be possible to obtain explicitly all derivatives $D^k\boldsymbol{\phi}(p')$ in terms of the derivatives $D^j\boldsymbol{\phi}(p')$, $j = 1, \ldots, k-1$, and partial derivatives of the functions $e_i(\boldsymbol{\phi})$, $i = 1, \ldots, m+1$, of degree up to $k+1$.

In order that the minimizing trajectory $\boldsymbol{\phi}(p')$ be expanded in a Taylor series about $p' = 0$, we have to show that limiting derivatives exist at $p' = 0$. For very large values of p, we can approximate the matrix as

$$\nabla(\nabla U(\boldsymbol{\phi}(p'), p))^{\mathrm{T}} \simeq q \left(\sum_{i \in K} \left(\frac{e_{ip'}}{M_{p'}} \right)^q \right)^{1/q-1} \sum_{i \in K} \left\{ \left(\frac{e_{ip'}}{M_{p'}} \right)^{q-2} \left(\frac{\nabla e_{ip'}}{M_{p'}} \right) (\nabla e_{ip'})^{\mathrm{T}} \right\}$$

$$= p H_p \tag{24}$$

where

$$H_p \triangleq (\operatorname{sgn} M_{p'}) M_{p'} s_q(p') \sum_{i \in K} \left\{ \frac{\mu_i(p')}{e_{ip'}{}^2} \nabla e_{ip'} (\nabla e_{ip'})^{\mathrm{T}} \right\} \tag{25}$$

and

$$s_q(p') \triangleq \left(\sum_{i \in K} \left(\frac{e_{ip'}}{M_{p'}} \right)^q \right)^{1/q} \tag{26}$$

$$\mu_i(p') \triangleq \frac{\left(\dfrac{e_{ip'}}{M_{p'}} \right)^q}{\displaystyle\sum_{i \in K} \left(\frac{e_{ip'}}{M_{p'}} \right)^q} \tag{27}$$

H_p is an $n \times n$ matrix and for any non-zero n-component vector $\mathbf{x}$:

$$\mathbf{x}^{\mathrm{T}} H_p \mathbf{x} = (\operatorname{sgn} M_{p'}) M_{p'} s_q(p') \sum_{i \in K} \left\{ \frac{\mu_i(p')}{e_{ip'}{}^2} \mathbf{x}^{\mathrm{T}} \nabla e_{ip'} (\nabla e_{ip'})^{\mathrm{T}} \mathbf{x} \right\} \tag{28}$$

Of interest is the positiveness of the terms $\mathbf{x}^{\mathrm{T}} \nabla e_{ip'} (\nabla e_{ip'})^{\mathrm{T}} \mathbf{x}$ in the summation. It follows that a necessary condition for $\mathbf{x}^{\mathrm{T}} H_p \mathbf{x}$ to be positive is that for the gradient vectors $\nabla e_{ip'}$, $i \in \check{K}$, at least n of them are linearly independent, where

$$\check{K} \triangleq \{ i \mid e_i(\boldsymbol{\phi}(0)) = M(\boldsymbol{\phi}(0)) \} \tag{29}$$

This ensures that the vector $\mathbf{x}$ cannot be orthogonal to the n gradient vectors $\nabla e_{ip'}$ simultaneously, and at least one of the terms $\mathbf{x}^{\mathrm{T}} \nabla e_{ip'} (\nabla e_{ip'})^{\mathrm{T}} \mathbf{x}$ will be positive. If the associated multipliers $\mu_i(p')$, $i \in \check{K}$, are positive, it is then sufficient for $\mathbf{x}^{\mathrm{T}} H_p \mathbf{x}$ to be positive and H_p be positive definite and hence invertible.

Therefore, (23) becomes

$$D\boldsymbol{\phi}(p') = (\operatorname{sgn} M_{p'})\{p H_p\}^{-1} q^2 \left(\sum_{i \in K} \left(\frac{e_{ip'}}{M_{p'}} \right)^q \right)^{1/q-1} \times \sum_{i \in K} \left\{ \left(\frac{e_{ip'}}{M_{p'}} \right)^{q-1} \ln \left(\frac{e_{ip'}}{M_{p'}} \right) \nabla e_{ip'} \right\}$$

$$= H_p^{-1} \left(\sum_{i \in K} \left(\frac{e_{ip'}}{M_{p'}} \right)^q \right)^{1/q-1} \sum_{i \in K} \left\{ \left(\frac{e_{ip'}}{M_{p'}} \right)^{q-1} \ln \left(\frac{e_{ip'}}{M_{p'}} \right)^q \nabla e_{ip'} \right\}$$

$$= H_p^{-1} \left(\sum_{i \in K} \left(\frac{e_{ip'}}{M_{p'}} \right)^q \right)^{1/q} \sum_{i \in K} \left\{ \frac{\left(\frac{e_{ip'}}{M_{p'}} \right)^q}{\sum_{i \in K} \left(\frac{e_{ip'}}{M_{p'}} \right)^q} \left[\ln \frac{\left(\frac{e_{ip'}}{M_{p'}} \right)^q}{\sum_{i \in K} \left(\frac{e_{ip'}}{M_{p'}} \right)^q} \right] \frac{M_{p'}}{e_{ip'}} \nabla e_{ip'} \right\}$$

$$= H_p^{-1} s_q(p') \sum_{i \in K} \left\{ \mu_i(p') [\ln \mu_i(p')] \frac{M_{p'}}{e_{ip'}} \nabla e_{ip'} \right\} \tag{30}$$

Imposing optimality conditions (Bandler and Charalambous 1973), we observe that

$$\lim_{p' \to 0} s_q(p') = 1 \tag{31}$$

$$\lim_{p' \to 0} \mu_i(p') = v_i \begin{cases} = 0, & i \notin \check{K} \\ \geqslant 0, & i \in \check{K} \end{cases} \tag{32}$$

$$\sum_{i \in \check{K}} v_i = 1 \tag{33}$$

$$\lim_{p' \to 0} \frac{e_{ip'}}{M_{p'}} = 1, \; i \in \check{K} \tag{34}$$

and the gradient vectors $\nabla e_i(\boldsymbol{\phi}(0))$, $i \in \check{K}$, are linearly dependent. Let us define $H_\infty = \lim_{p \to \infty} H_p$. Then, a necessary condition for H_∞ to be positive definite is that the set $\check{K}$ contains at least $n+1$ equal maxima and n of the associated gradient vectors $\nabla e_i(\boldsymbol{\phi}(0))$ are linearly independent. A sufficient condition is that the multipliers v_i, $i \in \check{K}$, are positive.

The limiting value of $D\boldsymbol{\phi}(p')$ at $p' = 0$ (or $p = \infty$) is therefore given by

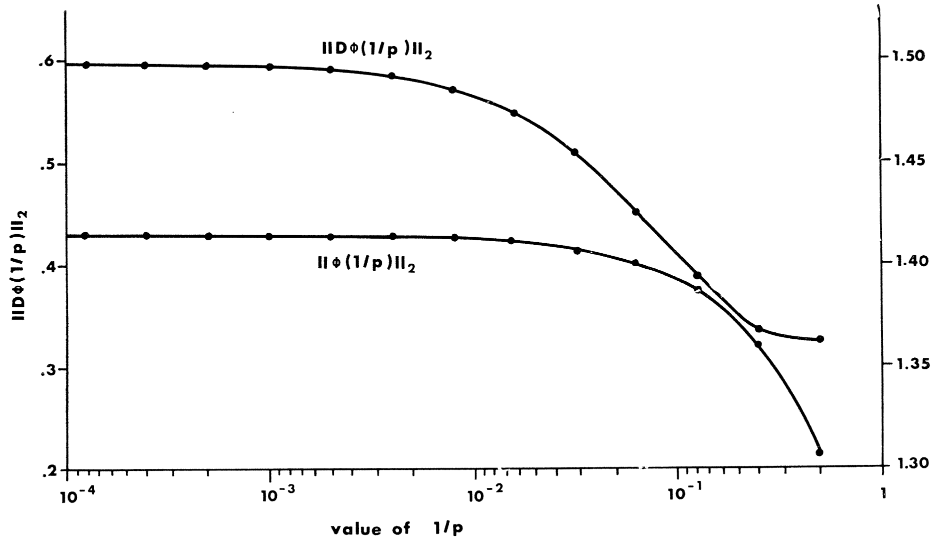
$$D\boldsymbol{\phi}(0) = \lim_{p' \to 0} D\boldsymbol{\phi}(p') = H_\infty^{-1} \sum_{i \in \check{K}} (v_i \ln v_i) \nabla e_i(\boldsymbol{\phi}(0)) \tag{35}$$

The existence of the higher-order derivatives of $\boldsymbol{\phi}(p')$ at $p' = 0$ may be derived in a similar manner.

To illustrate some of the ideas presented in this section, let us consider the minimization with respect to $\boldsymbol{\phi}$ of the maximum of the following three functions:

$$e_1(\boldsymbol{\phi}) = \phi_1^4 + \phi_2^2$$
$$e_2(\boldsymbol{\phi}) = (2 - \phi_1)^2 + (2 - \phi_2)^2$$
$$e_3(\boldsymbol{\phi}) = 2 \exp(-\phi_1 + \phi_2)$$

The minimax solution occurs at the point $\phi_1 = \phi_2 = 1$. The figure shows the Euclidean norms of $\boldsymbol{\phi}(1/p)$ and $D\boldsymbol{\phi}(1/p)$ as a function of $1/p$. The points correspond to minima for p in the sequence 2^i, $i = 1, \ldots, 14$. The limiting value of $||\boldsymbol{\phi}(1/p)||_2$ is $\sqrt{2}$ and $||D\boldsymbol{\phi}(1/p)||_2$ is well defined.



$||\boldsymbol{\phi}(1/p)||_2$ and $||D\boldsymbol{\phi}(1/p)||_2$ as a function of $1/p$.

4. Numerical results

A minimax example and two well-known test functions were used to illustrate the performance of the extrapolation formula (14) in estimating the minimax optimum from a sequence of least pth minima. A CDC 6400 computer was used throughout and Fletcher's (1972) recent quasi-Newton programme was used to perform the minimization. The initial estimate of the Hessian matrix (required in Fletcher's programme) was set to the unit matrix for the first optimization. In subsequent optimizations, the Hessian matrix computed at the previous minimum was used.

4.1. *Minimax example* (*Charalambous and Bandler* 1976)

For the minimax example of the previous section, using the least pth objective (1) and extrapolation formula (14) for $p = 4$, 16, 64, 256, 1024, 45 function evaluations yielded $\phi_1 = 1{\cdot}0000001$, $\phi_2 = 0{\cdot}9999999$. A less accurate solution of $\phi_1 = 1{\cdot}0000023$, $\phi_2 = 0{\cdot}9999945$ was obtained in 62 function evaluations using $p = 10^5$.

4.2. *Beale problem* (*Kowalik and Osborne* 1968)

Minimize

$$f(\boldsymbol{\phi}) = 9 - 8\phi_1 - 6\phi_2 - 4\phi_3 + 2\phi_1^2 + 2\phi_2^2 + \phi_3^2 + 2\phi_1\phi_2 + 2\phi_1\phi_3$$

subject to

$$\phi_i \geqslant 0, i = 1, 2, 3$$

$$3 - \phi_1 - \phi_2 - 2\phi_3 \geqslant 0$$

The function has a minimum $f(\check{\boldsymbol{\phi}})=1/9$ at $\check{\boldsymbol{\phi}}=[4/3\ 7/9\ 4/9]^{\mathrm{T}}$. The Bandler–Charalambous technique was used to transform the constrained problem into an unconstrained minimax problem. A sequence of least pth approximations together with extrapolation was used to obtain the optimal solution. The same problem was also solved by least pth approximation with a value of p of 10^5. The SUMT method of Fiacco and McCormick (1968) was also used to solve the problem by defining

$$U(\boldsymbol{\phi}, r)=f(\boldsymbol{\phi})-r \sum_{i=1}^{m} \ln g_i(\boldsymbol{\phi}) \tag{36}$$

and minimizing U w.r.t. $\boldsymbol{\phi}$ for a strictly decreasing sequence of r values together with extrapolation, also using the Fletcher programme under the same conditions. Table 1 gives a comparison between the three approaches.

	Least pth approach		Fiacco–McCormick method
Parameters	$p=4, 16, 64, 256$ $\alpha=1$ Order of extrapolation $=3$	$p=10^5$ $\alpha=1$	$r=10^{-2}, 2\times10^{-3}, 4\times10^{-4}, 8\times10^{-5}, 1{\cdot}6\times10^{-5}$ Order of extrapolation $=3$
ϕ_1	1·3333333	1·3333338	1·3333333
ϕ_2	0·7777778	0·7777775	0·7777778
ϕ_3	0·4444444	0·4444437	0·4444445
$f(\boldsymbol{\phi})$	0·1111111	0·1111114	0·1111111
$g_1(\boldsymbol{\phi})$	1·3333333	1·3333338	1·3333333
$g_2(\boldsymbol{\phi})$	0·7777778	0·7777775	0·7777778
$g_3(\boldsymbol{\phi})$	0·4444444	0·4444437	0·4444445
$g_4(\boldsymbol{\phi})$	$5{\cdot}07\times10^{-9}$	$1{\cdot}39\times10^{-6}$	$7{\cdot}82\times10^{-14}$
Function evaluations	34	78	40

Table 1. Results for the Beale problem for starting-point $\boldsymbol{\phi}^0=[1\ 2\ 1]^{\mathrm{T}}$.

4.3. *Rosen–Suzuki problem* (*Kowalik and Osborne* 1968)

Minimize

$$f(\boldsymbol{\phi})=\phi_1^2+\phi_2^2+2\phi_3^2+\phi_4^2-5\phi_1-5\phi_2-21\phi_3+7\phi_4$$

subject to

$$-\phi_1^2-\phi_2^2-\phi_3^2-\phi_4^2-\phi_1+\phi_2-\phi_3+\phi_4+8\geqslant 0$$

$$-\phi_1^2-2\phi_2^2-\phi_3^2-2\phi_4^2+\phi_1+\phi_4+10\geqslant 0$$

$$-2\phi_1^2-\phi_2^2-\phi_3^2-2\phi_1+\phi_2+\phi_4+5\geqslant 0$$

The function has a minimum $f(\check{\boldsymbol{\phi}})=-44$ at $\check{\boldsymbol{\phi}}=[0\ 1\ 2\ -1]^{\mathrm{T}}$. The Bandler–Charalambous technique was used to transform the non-linear programming problem into an unconstrained minimax problem. The minimax problem was then solved using a sequence of least pth approximations together with extrapolation and least pth approximation with a value of p of 10^5. The problem was also solved using the Fiacco–McCormick method with extrapolation with the same objective function of (36). Table 2 compares the performance of the three approaches.

4.4. *Comments*

In the three examples considered, the performance of the extrapolation procedure in yielding the solution of the minimax or non-linear programming problem is satisfactory. The order of estimates has been limited to three, though higher orders are possible. Computer storage requirements and accuracy considerations such as round-off error (which may become critical for higher-order estimates) prompted our choice. Numerical experience indicates that the factor c by which p_i is increased is not crucial to convergence. In general, the faster the rate of increase, the fewer are the number of minima required to obtain significant estimates of the solution values. Each minimum requires more computation to be reached than an increase at a slower rate. More minima are required to compute significant estimates in the latter case. A practical range for c is 2 to 10.

	Least pth approach		Fiacco–McCormick method
Parameters	$p=4, 12, 36, 108, 324, 972$ $\alpha=10$ Order of extrapolation $=3$	$p=10^5$ $\alpha=10$	$r=1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}$ Order of extrapolation $=3$
ϕ_1	$-0{\cdot}0000002$	$-0{\cdot}0000021$	$-0{\cdot}0000000$
ϕ_2	$1{\cdot}0000005$	$0{\cdot}9999976$	$1{\cdot}0000000$
ϕ_3	$1{\cdot}9999999$	$1{\cdot}9999908$	$2{\cdot}0000000$
ϕ_4	$-1{\cdot}0000002$	$-0{\cdot}9999883$	$-1{\cdot}0000000$
$f(\boldsymbol{\phi})$	$-44{\cdot}000000$	$-43{\cdot}999804$	$-44{\cdot}000000$
$g_1(\boldsymbol{\phi})$	$-2{\cdot}80\times10^{-7}$	$8{\cdot}56\times10^{-5}$	$-9{\cdot}35\times10^{-10}$
$g_2(\boldsymbol{\phi})$	$1{\cdot}00$	$1{\cdot}00$	$1{\cdot}00$
$g_3(\boldsymbol{\phi})$	$7{\cdot}57\times10^{-8}$	$5{\cdot}51\times10^{-5}$	$-7{\cdot}61\times10^{-11}$
Function evaluations	72	107	125

Table 2. Results for the Rosen–Suzuki problem for starting-point $\boldsymbol{\phi}^0=[0\ 0\ 0\ 0]^{\mathrm{T}}$.

5. Conclusions

Theoretical considerations and computational implications of applying an extrapolation technique in solving minimax and non-linear programming problems using a sequence of least pth approximations have been presented. Numerical results indicate that this approach is very promising. We note also that the least pth approach does not require a feasible starting-point, and that the efficiency depends mainly on the method used to determine the least pth minima.

References

Bandler, J. W., 1969, *I.E.E.E. Trans. microw. Theory Tech.*, **17**, 533.
Bandler, J. W., and Charalambous, C., 1972, *I.E.E.E. Trans. microw. Theory Tech.*, **20**, 834; 1973, *J. optimiz. Theory Applic.*, **11**, 556; 1974, *Ibid.*, **13**, 607.
Charalambous, C., and Bandler, J. W., 1976, *Int. J. Systems Sci.*, **7**, 377.
Fiacco, A. V., and McCormick, G. P., 1968, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques* (New York: Wiley).
Fletcher, R., 1972, *A.E.R.E. Rep.*, R7125, Atomic Energy Research Establishment, Harwell, Berkshire, England.
Joyce, D. C., 1971, *SIAM Rev.*, **13**, 435.
Kowalik, J., and Osborne, M. R., 1968, *Methods for Unconstrained Optimization Problems* (New York: Elsevier).
Lootsma, F. A., 1968, *Philips Res. Rep.*, **23**, 108.