MANAGEMENT SCIENCE
Vol. 22, No. 5, January, 1976
Printed in U.S.A.

LINE SEARCH TECHNIQUES BASED ON INTERPOLATING POLYNOMIALS USING FUNCTION VALUES ONLY*†

ARIE TAMIR‡
Northwestern University

In this study we derive the order of convergence of some line search techniques based on fitting polynomials, using function values only. It is shown that the order of convergence increases with the degree of the polynomial. If viewed as a sequence, the orders approach the Golden Section Ratio when the degree of the polynomial tends to infinity.

Introduction

Most efficient methods for unconstrained minimization utilize a one-dimensional search along directions generated by the method. If f is the function to be minimized, X the current vector of decision variables, and S the search direction, then the one-dimensional search problem is to choose $\alpha > 0$ yielding the first local minimum of $f(X + \alpha S)$. A significant portion of the total computational effort is expended in this search. The problem can be particularly difficult when f is an interior or exterior penalty function. This is a situation of great practical importance because penalty functions are widely used.

The most popular one-dimensional search procedures for use in unconstrained minimization utilize quadratic [2], [7] or 2 point cubic [1], [5], [7] interpolation of f. When applied to penalty functions these interpolation approaches have serious deficiencies. Quadratic interpolation has the drawback that its order of convergence is approximately 1.3, significantly less than that of 2 point cubic interpolation, which is 2 [8]. The 2-point cubic, however, requires the computation of ∇f. This is usually time consuming and is often difficult to code. In some cases ∇f may not be available analytically.

A one-dimensional search based on quadratic and cubic interpolations using functions values only is studied in [3]. The performance of this procedure on several test problems involving penalty functions has been significantly better than that of competing methods. The algorithm in [3] has motivated this study on order of convergence of related search techniques based on fitting polynomials, using functions values only.

The algorithm studied in this paper is as follows.

Let x be a scalar variable, and $f(x)$ the function to be minimized, assumed differentiable. An isolated minimum of f is assumed to occur at α, where

$$f'(\alpha) = 0. \tag{1}$$

Let n be a fixed integer greater than 1. If $x_i, x_{i-1}, \ldots, x_{i-n}$ are $n+1$ approximations to α, and $P_n(x)$ is the unique polynomial of degree less than or equal to n which satisfies

$$P_n(x_{i-j}) = f(x_{i-j}), \qquad j = 0, 1, \ldots, n, \tag{2}$$

then the new approximation to α, x_{i+1}, is chosen to satisfy

$$P_n'(x_{i+1}) = 0. \tag{3}$$

If $x_{i+1} \neq \alpha$ the procedure is repeated, fitting the next polynomial to $x_{i+1}, x_i, \ldots, x_{i-(n-1)}$. This algorithm is henceforth referred to as the Sequential Polynomial Fitting Algorithm (SPFA).

We note that the SPFA is different from the algorithm discussed in [3], in that the points through which the polynomial passes need not bracket a minimum of f. However, the bracketing algorithms do not lend themselves to the difference equation approach used in

* Processed by Professor Arthur M. Geoffrion, Departmental Editor for Integer and Large-Scale Programming; received April 4, 1974, revised November 28, 1974, March 19, 1975. This paper was with the author $4\frac{1}{2}$ months for revision.

† This research was partially supported by the Office of Naval Research under Contract No. N00014–67–A–0404–0010.

‡ The author is indebted to Professor L. S. Lasdon for his suggestions and assistance during the research work.

most convergence order derivations. Further, the two procedures are closely related. Other authors [4], [8] give intuitive arguments that the rate of convergence of sequential and bracketing algorithms is the same, then proceed to analyze the convergence order of the SPFA for the special case $n = 2$. We know of no proof that the convergence orders are the same, although the conjecture seems reasonable.

Convergence and Convergence Orders

In this work speed of convergence of line search methods is measured in terms of the following concepts. (See [8], [9].)

DEFINITION. Let the sequence $\{e_k\}$ converge to 0. The *order of convergence* of $\{e_k\}$ is defined as the supremum of the nonnegative numbers p satisfying $0 \leqslant \limsup_{k\to\infty}\{|e_{k+1}|/|e_k|^p\} < \infty$. (The case 0/0 is regarded as finite.) The *average order of convergence* is the infimum of the numbers $p > 1$ such that $\limsup_{k\to\infty}|e_k|^{1/p^k} = 1$. The order is infinity if the equality holds for no $p > 1$.

Let

$$J = \{x \mid |x - \alpha| \leqslant L\} \tag{4}$$

throughout this section, f is assumed to satisfy the following conditions. (The notation $f^{(i)}(x)$ denotes the ith derivative of f.)

Assumption 1. 1. $f^{(2)}(x) \neq 0$ for all $x \in J$. Note that this is equivalent to $f^{(2)}(x) > 0$ for all $x \in J$, since $f^{(2)}(\alpha) \neq 0$ and the minimality of α imply $f^{(2)}(\alpha) > 0$. 2. $f^{(n+1)}(\alpha) \neq 0$. 3. $f^{(n+2)}$ is continuous on J. 4. If we define constants M_0, M_1 and M_2 such that, for all $x \in J$,

$$|f^{(2)}(x)| \geqslant M_0, \quad |f^{(n+1)}(x)/(n+1)!| \leqslant M_1, \quad |f^{(n+2)}(x)/(n+2)!| \leqslant M_2 \tag{5}$$

then the interval width L in (4) is small enough to satisfy

$$\Gamma = L\left[2(M_2L + (n+1)M_1)/M_0\right]^{1/(n-1)} < 1 \quad \text{and} \tag{6}$$

$$\Gamma_1 = (1/M_0L)\left(M_2(2L)^{n+1} + M_1(n+1)(2L)^n\right) < \tfrac{1}{2}\,. \tag{7}$$

We note that if the constants M_0, M_1 and M_2 are defined as the sharpest possible bounds for a given L, then M_1 and M_2 are nondecreasing in L and M_0 is nonincreasing in L. Since Γ and Γ_1 are increasing functions of L, approaching zero as L approaches zero, an L satisfying (6) and (7) can always be found. Assumption 1 insures that the sequence $\{x_i - \alpha\}$ is well defined and converges to zero (Theorems 1 and 2).

We also require the assumption that the convergence rate of the sequence $\{e_i\}$ is at least of order 1, where

$$e_i = x_i - \alpha. \tag{8}$$

Assumption 2.[1]

$$\lim_{k\to\infty}\{e_{k+1}/e_k\} = \beta \tag{9}$$

where β is finite. In particular it is assumed throughout that the ratios e_{k+1}/e_k are well defined, i.e., $e_k \neq 0$ for all k. (Note, however, that $e_k = 0$ for some k implies that the SPFA converges in a finite number of steps.)

The main result of this section is

THEOREM 1. *Under Assumptions* 1 *and* 2, *the order of convergence of the SPFA*,

[1] The stronger version of this assumption, i.e., superlinear convergence, is made for the case $n = 2$ in [8, p. 143]. A discussion on the motivation and validity of Assumption 2 is provided in Appendix C. It is our conjecture that Assumption 2 is redundant and implied by Assumption 1.

using polynomials of degree n, is equal to the unique positive root, σ_n, of the polynomial

$$C_n(x) = x^{n+1} - \sum_{j=1}^{n} x^{n-j}. \tag{10}$$

The sequence of roots $\{\sigma_n\}$ is increasing, approaching the Golden Section Ratio $\tau = (1 + 5^{1/2})/2 \approx 1.618$ as n approaches infinity.

A table of positive roots of $C_n(x)$ is given below

n	root σ_n	σ_n/τ
2	1.324	0.81
3	1.465	0.90
4	1.534	0.94
5	1.570	0.97
6	1.590	0.98

Cubic polynomials ($n = 3$) yield 90% of the maximum attainable convergence order, and the ratio σ_n/τ increases slowly for $n > 3$. Given the added complexity of dealing with polynomials of degree greater than 3, there is little reason for considering such polynomials in practical interpolation schemes.

In the remainder of this section, we give a number of results leading to a proof of Theorem 1. The following two theorems, proved in Appendix A, insure that the sequence $\{x_i\}$ is well defined, and converges to the minimal point α. Note that Theorems 2, 3 are independent of Assumption 2 and applicable for all SPFA's.

THEOREM 2. *Define $J = \{x \mid |x - \alpha| \leqslant L\}$ and suppose that α is the unique minimum of f in J. Let $x_i, x_{i-1}, \ldots, x_{i-n}$ in J define the polynomial $P_n(x)$ of degree $\leqslant n$ satisfying $P_n(x_{i-j}) = f(x_{i-j})$, $j = 0, 1, 2, \ldots, n$. If f and J satisfy Assumption 1 then $P_n'(x)$ has a real root in J.*

THEOREM 3. *Suppose that the conditions of Theorem 2 hold and let x_{i+1} in J be a real root of the derivative of the interpolating polynomial $P_n(x)$ determined by $x_i, x_{i-1}, \ldots, x_{i-n}$. Then the sequence $\{x_k\}$ converges to α and*

$$|x_k - \alpha| \leqslant K\Gamma^{r(n,k)} \tag{11}$$

for some constant K where $\Gamma < 1$ (defined in (6)), and

$$r(n, k) = n^{k/(n+1)}. \tag{12}$$

Hence, for all SPFA's, the sequence $\{e_k\}$ converges to zero with average order of convergence greater than or equal to $n^{1/(n+1)}$.

We now derive results on the (stepwise) order of convergence of the SPFA. In Appendix A, it is shown that

$$P_n'(x) = f'(x) - \frac{f^{(n+1)}(\xi(x))}{(n+1)!} \sum_{k=0}^{n} \prod_{\substack{j=0;\\ j \neq k}}^{n} (x - x_{i-j}) - \frac{f^{(n+2)}(\eta(x))}{(n+2)!} \prod_{j=0}^{n} (x - x_{i-j}) \tag{13}$$

where $\xi(x)$ and $\eta(x)$ are in the interval determined by $x_i, x_{i-1}, \ldots, x_{i-n}, x$. Substituting $x = x_{i+1}$ into (13), and using the relations $P_n'(x_{i+1}) = 0$, $(x_{i+1} - x_{i-j}) = (e_{i+1} -$

$e_{i-j})$ and $f'(x_{i+1}) = e_{i+1} f^{(2)}(\theta(x_{i+1}))$ where $\theta(x_{i+1})$ is in the interval $[x_{i+1}, \alpha]$, yield

$$e_{i+1} f^{(2)}(\theta(x_{i+1})) = \frac{f^{(n+1)}(\xi(x_{i+1}))}{(n+1)!} \sum_{k=0}^{n} \prod_{\substack{j=0;\\ j \neq k}}^{n} (e_{i+1} - e_{i-j}) + \frac{f^{(n+2)}(\eta(x_{i+1}))}{(n+2)!} \prod_{j=0}^{n} (e_{i+1} - e_{i-j}). \tag{13a}$$

Suppose that $e_{i-j} \neq 0,\ j = 0, 1, \ldots, n$.

$$e_{i+1} f^{(2)}(\theta(x_{i+1})) = \prod_{j=1}^{n} e_{i-j} \left\{ \frac{f^{(n+1)}(\xi(x_{i+1}))}{(n+1)!} \left[\prod_{j=1}^{n} \left(\frac{e_{i+1}}{e_{i-j}} - 1 \right) + \sum_{k=1}^{n} \left(\frac{e_{i+1}}{e_{i-k}} - \frac{e_i}{e_{i-k}} \right) \prod_{\substack{j=1;\\ j \neq k}}^{n} \left(\frac{e_{i+1}}{e_{i-j}} - 1 \right) \right] + \frac{f^{(n+2)}(\eta(x_{i+1}))}{(n+2)!} (e_{i+1} - e_i) \prod_{j=1}^{n} \left(\frac{e_{i+1}}{e_{i-j}} - 1 \right) \right\}. \tag{14}$$

To prove results on the order of convergence we will require that Assumption 2 be satisfied. (See Appendix C.) We first derive lower and upper bounds on the order of convergence applying a relatively elementary result due to Ostrowski [10]. To strengthen the result and obtain the exact order of convergence we adopt an entirely different approach using more advanced results in the context of complex variable theory. Although the weak result is implied by the stronger conclusion, we present both approaches, since to our understanding, Ostrowski's result (Lemma 1), in spite of its elementary nature, is of great value in studies of convergence. Thus we start by demonstrating its applicability, hoping that the insight afforded compensates for the lack of brevity.

By Assumption 2 (i.e. $e_{i+1}/e_i \to \beta$), the ratios e_{i+1}/e_{i-j} in (14) approach β^{j+1} as $i \to \infty$, while $(e_{i+1} - e_i)$ approaches zero.

Defining A_{i+1} by

$$e_{i+1} = A_{i+1} \prod_{j=1}^{n} e_{i-j} \tag{15}$$

it is next proved that the linear convergence assumed above implies superlinear convergence (i.e. $\beta = 0$). Further, it is demonstrated that $A_{i+1} \to A \neq 0$. From (14)–(15) it follows that

$$A_{i+1} \to \Psi(\beta) f^{(n+1)}(\alpha)/(n+1)!\, f^2(\alpha) \equiv A \tag{16}$$

where

$$\Psi(x) = \prod_{j=1}^{n} (x^{j+1} - 1) + \sum_{k=1}^{n} (x^{k+1} - x^k) \prod_{\substack{j=1;\\ j \neq k}}^{n} (x^{j+1} - 1).$$

Suppose $\beta \neq 0$, then $e_{i-1}/e_i \to \beta^{-1}$ and (15) yields

$$\frac{e_{i+1}}{e_i} = A_{i+1} \frac{e_{i-1}}{e_i} \prod_{j=2}^{n} e_{i-j} \to 0.$$

Thus $\beta = 0$ and $\Psi(\beta) = \Psi(0) \neq 0$. By Assumption 1 (conditions 1, 2), $A \neq 0$. For i

sufficiently large and $\epsilon > 0$, (15) yields

$$(|A| - \epsilon) \prod_{j=1}^{n} |e_{i-j}| \leqslant |e_{i+1}| \leqslant (|A| + \epsilon) \prod_{j=1}^{n} |e_{i-j}| \tag{17}$$

or, defining

$$\delta_i = |e_i|(|A| + \epsilon)^{1/(n-1)}, \gamma_i = |e_i|(|A| - \epsilon)^{1/(n-1)}, \quad \delta_{i+1} \leqslant \prod_{j=1}^{n} \delta_{i-j}, \quad \gamma_{i+1} \geqslant \prod_{j=1}^{n} \gamma_{i-j}. \tag{18}$$

Since $e_i \neq 0$ for i sufficiently large, we may take logs of the inequalities (18) yielding the difference inequalities

$$\bar{c}_{i+1} \leqslant \sum_{j=1}^{n} \bar{c}_{i-j}, \quad c_{i+1} \geqslant \sum_{j=1}^{n} c_{i-j}, \quad \text{where} \tag{19}$$

$$\bar{c}_i = \ln \delta_i, \quad c_i = \ln \gamma_i. \tag{20}$$

We apply the following theorem, due to Ostrowski [10, p. 98], to (19).

LEMMA 1. *Consider* $x^n - \sum_{j=0}^{n-1} \rho_j x^j = 0$ *with* $\rho_j \geqslant 0$, $j = 0, \ldots, n-1$, *having positive root* σ, *and the infinite sequence* $\{u_i\}$ *satisfying the difference inequality*

$$u_{i+n} - \sum_{j=0}^{n-1} \rho_j u_{i+j} \geqslant 0, \qquad i = 1, 2, \ldots, \tag{21}$$

where $u_1, \ldots, u_n$ *are positive. Then we have* $u_i \geqslant \gamma\sigma^i$, $i = 1, 2, \ldots,$ *where* $\gamma = \min_{1 \leqslant j \leqslant n}(u_j/\sigma^j) > 0$.

By following Ostrowski's proof, it is easy to verify that, if the reverse inequality holds in (21) then $u_i \leqslant \delta\sigma^i$ where $\delta = \max_{1 \leqslant j \leqslant n}(u_j/\sigma^j) \geqslant \gamma$. Since $|e_i| \to 0$ we can assume without loss of generality that $\bar{c}_1, \ldots, \bar{c}_{n+1}$ and $c_1, \ldots, c_{n+1}$ are negative. Then, applying Lemma 1 to the sequences $\{-c_i\}$, $\{-\bar{c}_i\}$ yields $-\bar{c}_i \geqslant \gamma\sigma^i$, $-c_i \leqslant \delta\sigma^i$ where $\gamma = \min_{1 \leqslant j \leqslant n+1}(-\bar{c}_j/\sigma^j)$, $\delta = \max_{1 \leqslant j \leqslant n+1}(-c_j/\sigma^j)$, and σ is the unique positive root of the polynomial $C_n(x)$ in (10) as shown in Appendix B. Thus, using (20) and (18)

$$(|A| - \epsilon)^{-1/(n-1)} \exp(-\delta\sigma^i) \leqslant |e_i| \leqslant (|A| + \epsilon)^{-1/(n-1)} \exp(-\gamma\sigma^i). \tag{22}$$

Hence

$$|e_{i+1}|/|e_i|^t \leqslant g_1 \exp\{\sigma^i(\delta t - \gamma\sigma)\}, \quad \text{where} \quad g_1 < \infty. \tag{23}$$

In Appendix B we show that $\sigma > 1$. Hence the right-hand side of (23) is finite for all i if $(\delta t - \gamma\sigma) \leqslant 0$, i.e., if $t \leqslant \gamma\sigma/\delta$. Again using (22) we obtain

$$|e_{i+1}|/|e_i|^t \geqslant g_2 \exp\{\sigma^i(\gamma t - \delta\sigma)\}, \qquad g_2 > 0,$$

which approaches infinity as $i \to \infty$ if $\gamma t - \delta\sigma > 0$, i.e., if $t > \delta\sigma/\gamma$. Hence the order of convergence of the SPFA is less than or equal to $\delta\sigma/\gamma$ and greater than or equal to $\gamma\sigma/\delta$.

To show that the order of convergence is exactly σ, we use the following lemma [10, p. 92].

LEMMA 2. *Consider the linear difference equation* $u_{i+1} = k_{i+1} + \sum_{j=0}^{n} a_j u_{i-j}$, $i = n, n+1, \ldots,$ *where the* a_j *are constants and* $\{k_i\}$ *is a specified sequence. The associated characteristic polynomial is* $Q(x) = x^{n+1} - \sum_{j=0}^{n} a_j x^{n-j}$. *Let* $r_1, \ldots, r_{n+1}$ *be the roots of* $Q(x)$, *with* $|r_1| \geqslant |r_2| \geqslant \cdots \geqslant |r_{n+1}|$. *Assume that* $|r_1| > 1 > |r_2|$ *and, for* *some* s, $0 < s < |r_1|$, $k_i = O(s^i)$ *which means* $|k_i|/s^i \to c$ *for some constant* c *as* $i \to \infty$. *Then there exists* α_1 *such that, as* $i \to \infty$, $u_i/r_1^i \to \alpha_1$. *In addition, if* $s > |r_2|$, $u_i = \alpha_1 r_1^i + O(s^i)$. *If* $s = |r_2|$ *and* m *is the maximum multiplicity of all zeros of* $Q(x)$ *with modulus* $|r_2|$ *then* $u_i = \alpha_1 r_1^i + O(i^m |r_2|^i)$.

A careful examination of the proof in [10] shows that Lemma 2 is true even if the condition $|r_1| > 1 > |r_2|$ is replaced by the weaker condition $|r_1| > 1$, $|r_1| > |r_2|$. Taking absolute values and logs of (15), and defining $d_i = \ln|e_i|$, $B_i = \ln|A_i|$ we obtain $d_{i+1} = B_{i+1} + \sum_{j=1}^{n} d_{i-j}$, $i = n, n+1, \ldots$. Further defining $u_i = d_i/(\ln|A| + S)$, $k_i = B_i/(\ln|A| + S)$ where $S = -1$ if $|A| < 1$ and $S = 1$ otherwise, yields

$$u_{i+1} = k_{i+1} + \sum_{j=1}^{n} u_{i-j}, \qquad i = n, n+1, \ldots, \tag{24}$$

where, for i sufficiently large

$$|k_{i+1}| < 1. \tag{25}$$

The characteristic polynomial of (24) is $C_n(x)$ in (10). Consider first the case where $n+1$ is odd. It is shown in Appendix B that, in this case, the roots of $C_n(x)$ satisfy $|r_1| > 1 > |r_2|$. By (25), we can apply Lemma 2 with $s = 1$ to obtain $u_i = \alpha_1 r_1^i + O(1)$ implying $|e_i| = \exp\{-\beta_1 r_1^i + O(1)\}$ where $\beta_1 > 0$ since $|e_i| \to 0$. This implies that

$$|e_{i+1}|/|e_i|^t = \exp\{\beta_1 r_1^i (t - r_1) + O_1(1) + tO_2(1)\}$$

which implies that the order of convergence of the sequence $\{e_i\}$ is r_1. Suppose now that $n+1$ is even. Then, from Appendix B, $r_1 > 1$ and $r_2 = -1$. The comment following Lemma 2 justifies its use in this circumstance and, using $s = |r_2| = 1$ we obtain $u_i = \alpha_1 r_1^i + O(i^m |r_2|^i)$. As shown in Appendix B, $m = 1$, so $u_i = \alpha_1 r_1^i + O(i)$ which implies

$$|e_i| = \exp\{\gamma_1 r_1^i + O(i)\}. \tag{26}$$

Since $|e_i| \to 0$, $\gamma_1 \leqslant 0$. If $\gamma_1 = 0$ then $|e_i| = \exp\{O(i)\}$, which contradicts (11). Hence $\gamma_1 < 0$. It is then easily verified that (26) implies that the order of convergence of the sequence $\{e_i\}$ is again r_1. Theorem 1 follows from the preceding discussion and Appendix B.

A concluding remark is in order. The above discussion depends substantially on the assumption that $f^{(2)}(\alpha) \neq 0$. In fact we can weaken this assumption as follows. Suppose that $f^{(r)}(\alpha) = 0$, $r = 1, \ldots, k-1$, and $f^{(k)}(\alpha) \neq 0$, where $n + 1 > k \geqslant 2$. The minimality of α implies that k is even and $f^{(k)}(\alpha) > 0$. It is easy to verify that Theorem 2 is still valid if M_0 is the minimum of $f^{(k)}(\alpha)$ on J and (7) is replaced by

$$[(k-1)!/M_0 L^{k-1}]\left(M_2 (2L)^{n+1} + M_1 (n+1)(2L)^n\right) < \tfrac{1}{2}.$$

Theorem 1 is also valid if σ_n is replaced by θ_n where $\theta_n > 1$ and is the unique positive root of the polynomial $(k-1)x^{n+1} - x^{n-1} - x^{n-2} - \cdots - 1$. The sequence of roots $\{\theta_n\}$ is increasing and converges to $\theta = \{1 + [1 + 4/(k-1)]^{1/2}\}/2$.

We also note that τ, the bound on the convergence rates of interpolating polynomials using function values only, is easily exceeded when derivative values are incorporated to define the polynomial. For example, the quadratic obtained using the False Position method converges with rate equal to the Golden Section Ratio τ (see [8]). This method utilizes values of the function and its first derivative only. Newton's method uses second derivatives and has rate equal to 2.

Appendix A. Existence Theorem of a Zero of the Derivative of the Interpolation Polynomial

In this appendix we prove Theorems 2 and 3, assuring that the sequence of roots $\{x_i\}$, generated by the algorithm, is well defined in the neighborhood of α, and converges to α.

PROOF OF THEOREM 2. Since $f^{(n+1)}(x)$ is continuous, it is well known (e.g. [12, p. 61]) that

$$f(x) = P_n(x) + \frac{f^{(n+1)}(\xi(x))}{(n+1)!} \prod_{j=0}^{n} (x - x_{i-j}) \tag{A.1}$$

where $\xi(x)$ lies in the interval determined by $x_i, x_{i-1}, \ldots, x_{i-n}, x$. To derive an expression for $P_n'(x)$ we apply a result due to Ralston [11], which states that

$$\frac{1}{(n+1)!} \frac{d}{dx} f^{(n+1)}(\xi(x)) = \frac{1}{(n+2)!} f^{(n+2)}(\eta(x)) \tag{A.2}$$

where $\eta(x)$ is again a mean value in the interval of interpolation. Differentiating (A.1) and using (A.2) yield

$$P_n'(x) = f'(x) - \frac{f^{(n+1)}(\xi(x))}{(n+1)!} \sum_{k=0}^{n} \prod_{\substack{j=0;\\ j\neq k}}^{n} (x - x_{i-j}) - \frac{f^{(n+2)}(\eta(x))}{(n+2)!} \prod_{j=0}^{n} (x - x_{i-j}). \tag{A.3}$$

We now show that under the assumptions of the theorem $P_n'(x)$ has a zero in J. Note first that $f^{(2)}(x) > 0$ $\forall x \in J$ since α is a minimum point and hence $f^{(2)}(\alpha) \geqslant 0$. The theorem follows when we prove that $P_n'(\alpha - L) < 0$ and $P_n'(\alpha + L) > 0$. $f'(\alpha) = 0$ implies $f'(x) = f'(x) - f'(\alpha) = (x - \alpha) f^{(2)}(\gamma(x))$ where $\gamma(x)$ is in J. Substituting $x = \alpha - L$ in (A.3) yields

$$P_n'(\alpha - L) = -Lf^{(2)}(\gamma(\alpha - L)) - \frac{f^{(n+1)}(\xi(\alpha - L))}{(n+1)!} \sum_{k=0}^{n} \prod_{\substack{j=0;\\ j\neq k}}^{n} (\alpha - L - x_{i-j})$$

$$- \frac{f^{(n+2)}(\eta(\alpha - L))}{(n+2)!} \prod_{j=0}^{n} (\alpha - L - x_{i-j}).$$

$P_n'(\alpha - L)$ is negative if

$$T = \frac{1}{Lf^{(2)}(\gamma(\alpha - L))} \left[-\frac{f^{(n+1)}(\xi(\alpha - L))}{(n+1)!} \sum_{k=0}^{n} \prod_{j\neq k} (\alpha - L - x_{i-j}) - \frac{f^{(n+2)}(\eta(\alpha - L))}{(n+2)!} \prod_{j=0}^{n} (\alpha - L - x_{i-j}) \right] < 1.$$

But $T \leqslant |T| \leqslant (M_2/M_0)((2L)^{n+1}/L) + (M_1(n+1)/M_0)((2L)^n/L) < 1$. Similar arguments lead to the conclusion that $P_n'(\alpha + L) > 0$, and hence the theorem follows.

PROOF OF THEOREM 3. Substituting $x = x_{i+1}$ in (A.3) we obtain

$$f'(x_{i+1}) = \frac{f^{(n+1)}(\theta_1)}{(n+1)!} \sum_{l=0}^{n} \prod_{\substack{j=0;\\ j\neq l}}^{n} (x_{i+1} - x_{i-j}) + \frac{f^{(n+2)}(\theta_2)}{(n+2)!} \prod_{j=0}^{n} (x_{i+1} - x_{i-j})$$

where $\theta_1 = \xi(x_{i+1})$, $\theta_2 = \eta(x_{i+1})$. Defining $e_k = x_k - \alpha$, $k = 1, 2, \ldots,$ and noting that $f'(x_{i+1}) = e_{i+1} f^{(2)}(\theta_3)$, $\theta_3 = \gamma(x_{i+1})$ yield

$$M_0 |e_{i+1}| \leqslant M_1 \sum_{l=0}^{n} \left[|e_{i+1}| L^{n-1}(2^n - 1) + \prod_{j\neq l} |e_{i-j}| \right] + M_2 \left[|e_{i+1}| L^n (2^{n+1} - 1) + \prod_{j=0}^{n} |e_{i-j}| \right].$$

Hence,

$$|e_{i+1}| \leqslant \left\{ \frac{M_1}{M_0} \frac{(2L)^n}{L} (n+1) + \frac{M_2}{M_0} \frac{(2L)}{L}^{n+1} \right\} |e_{i+1}|$$

$$+ \frac{M_1}{M_0} (n+1) \operatorname*{Max}_{0 \leqslant j \leqslant n} |e_{i-j}|^n + \frac{M_2}{M_0} \operatorname*{Max}_{0 \leqslant j \leqslant n} |e_{i-j}|^{n+1}.$$

By Assumption 1, $(M_1/M_0)((2L)^n(n+1)/L) + (M_2/M_0)((2L)^{n+1}/L) < \frac{1}{2}$. Thus,

$$|e_{i+1}| \leqslant C \operatorname*{Max}_{0 \leqslant j \leqslant n} |e_{i-j}|^n \tag{A.5}$$

where $C = 2(M_1(n+1)/M_0 + M_2L/M_0)$. Define $\bar{e}_{i+1} = |e_{i+1}|C^{1/(n-1)}$. Then (A.5) yields

$$\bar{e}_{i+1} \leqslant \max_{0 \leqslant j \leqslant n} \bar{e}^n_{i-j}. \tag{A.6}$$

We show that if $k = t(n+1) + l$, $t \geqslant 1$, $l = 0, 1, \ldots, n$, then (A.6) implies that $\bar{e}_k \leqslant \Gamma^{n^t}$ where $\Gamma = LC^{1/(n-1)}$. The proof is by induction on k.

Let $t = 1$, $l = 0$ and consider $\bar{e}_{n+1}$, then $\bar{e}_{n+1} \leqslant \text{Max}\{\bar{e}^n_0, \bar{e}^n_1, \ldots, \bar{e}^n_n\} \leqslant \Gamma^n$. Let $k = t(n+1) + l$ and suppose that the result holds for indices smaller than k. If $l = 0$, then

$$\bar{e}_k = \bar{e}_{t(n+1)} \leqslant \max_{0 \leqslant l \leqslant n} \left[\bar{e}_{(t-1)(n+1)+l}\right]^n \leqslant \Gamma^{n^{t-1}n} = \Gamma^{n^t}.$$

Let $l > 1$, then

$$\bar{e}_k = \bar{e}_{t(n+1)+l} \leqslant \text{Max}\left\{ \max_{0 \leqslant j < l} [\bar{e}^n_{t(n+1)+j}],\ \max_{l \leqslant j \leqslant n} [\bar{e}^n_{(t-1)(n+1)+j}] \right\} \leqslant \text{Max}\{\Gamma^{n^t}, \Gamma^{n^{t+1}}\} \leqslant \Gamma^{n^t}.$$

Hence $\bar{e}_k \leqslant \Gamma^{n^t}$ where $k = t(n+1) + l$, $t \geqslant 1$, $l = 0, 1, \ldots, n$, so $|e_k| = \bar{e}_k C^{-1/(n-1)} \leqslant C^{-1/(n-1)}\Gamma^{n^t}$. $t = k/(n+1) - l/(n+1)$ and $\Gamma < 1$ (Assumption 1) imply that $|e_k| \leqslant C^{-1/(n-1)}\Gamma^{r(n,k)}$ and the theorem follows.

Appendix B. The Roots of the Indicial Equation

In this appendix we study the properties and roots of the polynomial

$$C_{k-1}(z) = z^k - z^{k-2} - z^{k-3} - \cdots - 1. \tag{B.1}$$

We will show that $C_{k-1}(z)$, $k \geqslant 3$, has a unique simple positive root, σ_{k-1}, with modulus greater than 1, and that all other roots are also simple with moduli less than or equal to 1. In fact, it will be proved that if k is odd σ_{k-1} is the only real root and that the other $k-1$ roots are inside the unit disc. If k is even $z = -1$ and σ_{k-1} are the only real roots and the other $k-2$ roots have moduli less than 1. It is also shown that the sequence $\{\sigma_k\}$, $k = 2, 3, \ldots$, is increasing and tends to the Golden Section ratio, τ (i.e., $\tau^2 - \tau - 1 = 0$, $\tau > 1$).

LEMMA B.1. *Let $C_{k-1}(z)$, $k \geqslant 3$, be defined by* (B.1). *$C_{k-1}(z)$ has a unique simple positive root, σ_{k-1}, and $1 < \sigma_{k-1} < \tau$, where τ is the Golden Section ratio. If k is odd σ_{k-1} is the only real root, and if k is even $z = -1$ is the only other real root of $C_{k-1}(z)$ and is simple.*

PROOF. $C_{k-1}(z) = z^k - (z^{k-1} - 1)/(z-1) = [z^{k-1}(z^2 - z - 1)]/(z-1) + 1/(z-1)$. Let τ be the Golden Section ratio, i.e., $\tau^2 - \tau - 1 = 0$, $\tau > 1$. $C_{k-1}(\tau) = \tau^{k-1}(\tau^2 - \tau - 1)/(\tau - 1) + 1/(\tau - 1) = 1/(\tau - 1) > 0$. It is easy to verify that for $k \geqslant 3$, $C_{k-1}(1) < 0$, and hence there exists a positive root $1 < \sigma_{k-1} < \tau$. To see that σ_{k-1} is simple and also the unique positive root observe first that

$$C_{k-1}(z) = (z - \sigma_{k-1})(z^{k-1} + a_2 z^{k-2} + a_3 z^{k-3} + \cdots + a_k z^0),$$

$$a_2 = \sigma_{k-1}, \quad a_i = (a_{i+1} + 1)/\sigma_{k-1}, \quad k > i \geqslant 3, \quad a_k = 1/\sigma_{k-1}.$$

Thus $a_i > 0$, $i = 2, 3, \ldots, k$, and the result follows. Suppose that k is even, then $C_{k-1}(z) = (z+1)(z^{k-1} - z^{k-2} - z^{k-4} - \cdots - 1)$. Hence $z = -1$ is a simple root. It is easily verified that $C_{k-1}(t)$ is negative for $-1 < t \leqslant 0$ and positive for $t < -1$. Hence $z = -1$ is the unique nonpositive real root. Suppose now that $k \geqslant 3$ is odd:

$$C_{k-1}(z) = z^k - (z+1)(z^{k-3} + z^{k-5} + \cdots + 1) = z^{k-1}(z^2 - z - 1)/(z-1) + 1/(z-1).$$

From the first expression we see that $C_{k-1}(t) < 0$ for $-1 \leqslant t \leqslant 0$. Now let $t < -1$. Clearly $t^2 - t - 1 > 0$ and $C_{k-1}(t) < 0$. Hence $C_{k-1}(t) < 0$ for all nonpositive t and σ_{k-1} is the unique real root.

LEMMA B.2. *All the roots of $C_{k-1}(z)$ are simple.*

PROOF. Define $D_{k-1}(z) = (z-1)C_{k-1}(z) = z^k(z-1) - z^{k-1} + 1$. If $z \neq 1$ is a multiple root of $C_{k-1}(z)$ it is also a multiple root of $D_{k-1}(z)$ and $D'_{k-1}(z) = 0$, $(k+1)z^k - kz^{k-1} - (k-1)z^{k-2} = 0$, $z = 0$ is not a root and we have $(k+1)z^2 - kz - (k-1) = 0$ which implies that z is real. The preceding lemma assures that real roots are simple and the lemma follows.

The following lemma shows that the sequence $\{\sigma_k\}$ is an increasing one.

LEMMA B.3. *$\{\sigma_k\}$, $k = 2, 3, \ldots$, is an increasing sequence and* $\lim_k \sigma_k = \tau$.

PROOF. To show the monotonicity property we prove that $C_k(\sigma_{k-1}) < 0$. Lemma B.1 then assures that $\sigma_k > \sigma_{k-1}$.

$$C_k(z) = \frac{z^{k+1}(z-1) - z^k + 1}{z-1} = z\left(C_{k-1}(z) - \frac{1}{z-1}\right) + \frac{1}{z-1},$$

$$C_k(\sigma_{k-1}) = \sigma_{k-1}\left(0 - \frac{1}{\sigma_{k-1}-1}\right) + \frac{1}{\sigma_{k-1}-1} = -1.$$

The sequence $\{\sigma_k\}$ is a bounded increasing sequence and hence $\lim_k \sigma_k = \beta$ exists. $\sigma_{k-1}^{k-1}(\sigma_{k-1}^2 - \sigma_{k-1} - 1) = -1$, $1 < \sigma_k < \tau$, $\Rightarrow \beta^2 - \beta - 1 = 0$ and $\beta = \tau$, the Golden Section root.

To prove that the $(k-1)$ roots of $C_{k-1}(z)$ that differ from σ_{k-1} have moduli less than or equal to 1, we introduce the following two results.

THEOREM B.1 (TRAUB [12, p. 51]). *Let $f_k(z) = z^k - a(z^{k-1} + z^{k-2} + \cdots + 1)$, $ka > 1$ and $k \geq 2$. Then $f_k(z)$ has one positive simple root, γ_k, and $\max(1, a) < \gamma_k < 1 + a$. All other roots are also simple with moduli less than* 1.

LEMMA B.4 (OSTROWSKI [10, p. 222]). *Let B be a closed region in the z-plane, the boundary of which consists of a finite number of regular arcs, and let $f(z)$ and $h(z)$ be regular on B. Assume that for no value of the real parameter t, running through the interval $a \leq t \leq b$, the function $f(z) + th(z)$ becomes zero on the boundary of B. Then the number $N(t)$ of the zeroes of $f(z) + th(z)$ inside B is independent of t for $a \leq t \leq b$.*

We are now ready to prove the main result.

THEOREM B.2. *If k is odd the $k-1$ roots of $C_{k-1}(z)/(z - \sigma_{k-1})$ have moduli less than* 1. *If k is even the $k-2$ roots of $c_{k-1}(z)/(z - \sigma_{k-1})(z+1)$ have moduli less than* 1.

PROOF. Let $\epsilon > 0$ be arbitrarily small and $k \geq 3$ and consider the polynomial $C_{k-1}(z) - tz^{k-1}$ for $t \in [\epsilon, 1]$ where $C_{k-1}(z) = z^k - z^{k-2} - z^{k-3} - \cdots - 1$. We show that $C_{k-1}(z) - tz^{k-1} \neq 0$ for all z in $\{z \mid |z| = 1\}$. Since $C_{k-1}(1) - t < 0$ $\forall t \in [\epsilon, 1]$ it is sufficient to show that $(z-1)\{C_{k-1}(z) - tz^{k-1}\} \neq 0$ for all $z \neq 1$ and $|z| = 1$. Suppose $(z-1)\{C_{k-1}(z) - tz^{k-1}\} = 0$ for some $z \neq 1$ and $|z| = 1$. Then

$$\{z^{k-1}[z^2 - z(t+1) - (1-t)] + 1\} = 0,$$

$$|z^{k-1}|\,|z^2 - z(t+1) - (1-t)| = |-1| \Rightarrow |z^2 - z(t+1) - (1-t)| = 1.$$

If $z = \cos\theta + i \sin$, then $[\cos 2\theta - (t+1)\cos\theta - (1-t)]^2 + [\sin 2\theta - (t+1)\sin\theta]^2 = 1$, which yields $-2(1-t)\cos^2\theta - (t+t^2)\cos\theta + (1+t^2) + (1-t) = 0$. Let $y = \cos\theta$ then it is clear that $y = 1$ is one root of the quadratic

$$2(1-t)y^2 + (t+t^2)y - (t^2 - t + 2) = 0. \tag{B.2}$$

For $t = 1, y = 1$ is the only root and we obtain $\cos\theta = 1$ which contradicts the assumption $z \neq 1$. Let $t \in [\epsilon, 1)$, then the second root of (B.2) is $y(t) = -(t^2 - t + 2)/2(1-t)$, $y(t) < (2t-2)/2(1-t) = -1$. Thus we have the contradiction $\cos\theta < -1$. Observing that for $t = 1$, $C_{k-1}(z) - tz^{k-1}$ yields the polynomial $f_k(z)$ with $a = 1$, discussed in Theorem B.1, we apply Lemma B.4 to conclude that for any positive t arbitrarily close to zero the polynomial $C_{k-1}(z) - tz^{k-1}$ has $k-1$ roots *inside* the disc $\{z \mid |z| \leq 1\}$. Continuity arguments (see for example [10, Appx. A]) lead to the conclusion that $C_{k-1}(z)$ has $k-1$ roots in $\{z \mid |z| \leq 1\}$. By substituting $t = 0$ in (B.2) we easily verify that the only possible root of $C_{k-1}(z)$ on the boundary of the disc is $z = -1$ which is a root if and only if k is even. Hence, the theorem is proved.

Appendix C. On the Linear Convergence Assumption

In Theorems 2 and 3 it is shown that the error sequence $\{e_i\}$ is well defined and converges (with average order of at least $n^{1/(n+1)}$) to zero, provided the initial estimates of α, the zero of $f'(x)$, are close enough to α (Assumption 1). To obtain results on the (stepwise) order of convergence we have assumed that (Assumption 2) $\{e_i\}$ is "well behaved", i.e., e_{i+1}/e_i tends to a finite limit.

Although we cannot provide a rigorous proof, covering an arbitrary distribution of the initial estimates of α in a small enough locality, it is our conjecture that Assumption 2 is implied by the conditions of Assumption 1. To strengthen and motivate the above statement we appeal to the result of Theorem 3 that applies to all SPFA's satisfying Assumption 1 only. We sketch a proof of the validity of the conjecture for some choice of initial estimates.

It is shown in Appendix A (A.5) that

$$|e_{i+1}| \leqslant C \max_{0 \leqslant j \leqslant n} |e_{i-j}|^n \tag{C.1}$$

where $C = 2(M_1(n+1)/M_0 + M_2L/M_0)$. In fact (C.1) can be improved as follows. Using (13a) we define $N_i, M_i^0, M_i^1, \ldots, M_i^n$ by

$$N_i e_{i+1} = \sum_{k=0}^{n} M_i^k \prod_{\substack{j=0;\\ j \neq k}}^{n} e_{i-j} \tag{C.2}$$

where $N_i \to_i f^{(2)}(\alpha)$ and $M_i^k \to_i f^{(n+1)}(\alpha)/(n+1)!$ for $k = 0, 1, \ldots, n$.

From Assumption 1 (conditions 1, 2) it follows that these sequences of coefficients converge to nonzero limits. Thus the asymptotic behavior of e_{i+1} is determined by that of $\sum_{k=0}^{n} \prod_{j=0; j \neq k}^{n} e_{i-j}$. Hence, for the sake of this (heuristic) discussion we assume that for i sufficiently large:

$$e_{i+1} = A \sum_{k=0}^{n} \prod_{\substack{j=0;\\ j \neq k}}^{n} e_{i-j} \quad \text{where} \quad A \neq 0. \tag{C.3}$$

To simplify the discussion, let us further assume that $A = 1$ (otherwise a normalization procedure similar to (A.5)—(A.6) would have been applied).

To motivate our conjecture we show that if (C.3) holds (with $A = 1$), and if the initial $(n+1)$ estimates are on one side of α and monotonically improving (i.e., $e_k > e_{k+1} > 0$, $k = 0, 1, \ldots, n-1$) then e_{i+1}/e_i tends to a finite limit. We comment that the above assumption is not (intuitively) very optimistic since the initial estimates do not bracket the minimum α.

We first note that it is sufficient to demonstrate that $\{e_i\}$ is a decreasing sequence since then (C.3) (with $A = 1$) implies $e_{i+1} = \prod_{j=1}^{n} e_{i-j}(1 + e_i/e_{i-1} + \cdots + e_i/e_{i-n}) = R_{i+1}\prod_{j=1}^{n} e_{i-j}$, where $1 \leqslant R_{i+1} \leqslant n+1$. As shown in the main text (by applying Lemma 2) the above difference equation with the boundedness of R_{i+1} suffices for e_{i+1}/e_i^t to converge when $t \geqslant 1$ is the only positive root of (10).

It is easy to verify that if e_k, $k = 0, \ldots, n$, are small enough (e.g. less than $(n+1)^{-1}$), then $0 < e_{n+1} \leqslant e_k$, $k = 0, 1, \ldots, n-1$. If also $e_{n+1} \leqslant e_n$, then $e_0, e_1, \ldots, e_{n+1}$ is a decreasing sequence of $n+2$ consecutive iterates. It is then implied by (C.3) that $\{e_i\}_{i=0}^{\infty}$ is a decreasing sequence and the result follows. Hence suppose $e_{n+1} > e_n$. (C.3) then yields $e_{n+k} > e_{n+k+1}$, $k = 1, 2, \ldots, n$. Repeating this argument it can then be proved by induction that for all integer $t \geqslant 0$,

$$0 < e_{i+1} < e_i, \quad tn + t \leqslant i < (t+1)n + t \tag{C.4}$$

and $e_{tn+t} \leqslant e_{tn+t-2}$ for $t \geqslant 1$. To complete the proof we will show that there exists a decreasing sequence of $(n+2)$ consecutive iterates. By the above discussion it clearly suffices to demonstrate that for some t

$$e_{(t+1)n+(t+1)} < e_{(t+1)n+t} \tag{C.5}$$

The sequence $\{e_i\}$ converges to zero, hence,

$$e_{(t+1)n+t} \underset{t}{\longrightarrow} 0. \tag{C.6}$$

Thus, there exists t_0 such that

$$e_{(t_0+1)n+t_0} > e_{(t_0+2)n+(t_0+1)}. \tag{C.7}$$

Using (C.3) for $i+1$, $i+2$ and subtracting yield

$$e_{i+2} - e_{i+1} = (e_{i+1} - e_{i-n}) \sum_{k=0}^{n-1} \prod_{\substack{j=0;\\ j \neq k}}^{n-1} e_{i-j}. \tag{C.8}$$

Choosing $i = (t_0+2)n + t_0$ to have $i - n = (t_0+1)n + t_0$, (C.7) and (C.8) yield $e_{(t_0+2)n+(t_0+2)} - e_{(t_0+2)n+(t_0+1)} < 0$ which proves (C.5) for $t = t_0 + 1$. Thus the proof is complete.

Assuming (C.3) we have proved that for sets of $(n+1)$ initial estimates $e_0, \ldots, e_n$

satisfying $0 < e_{k+1} < e_k$, $k = 0, \ldots, n-1$, the generated sequence of error terms $\{e_i\}$ is decreasing for i sufficiently large and hence Assumption 2 is satisfied.

A final comment is in order. It seems that a rigorous proof of the conjecture, along the lines suggested above, will require an argument justifying the substitution of A for the sequences M_i^k/N_i in (C.2) to yield (C.3).

We note that such substitutions for linear difference equations are discussed in [8], [10]. Also note that in the derivation of the order of convergence we in fact applied Lemma 2 to allow the substitution of A_{i+1} by A in (15). Unlike (15) which can readily be transformed to a nonhomogeneous linear difference equation, (C.3) involves sums of nonlinear terms and hence a different approach may be needed.

References

1. Fletcher, R. and Powell, M. J. D., "A Rapidly Convergent Descent Method for Minimization," *Brit. Computer J.*, Vol. 6 (1963), pp. 163–168.
2. ——— and Reeves, C. M, "Function Minimization by Conjugate Gradients," *Brit. Computer J.*, Vol. 7 (1964), pp. 149–154.
3. Fox, R. L., Lasdon, L. S., Tamir, A. and Ratner, M. W., "An Efficient One-Dimensional Search Procedure," *Management Science* (to appear).
4. Kowalik, J. and Osborne, M. R., *Methods for Unconstrained Optimization Problems*, Elsevier, New York, 1968.
5. Lasdon, L. S., Fox, R. and Ratner, M., "An Efficient One-Dimensional Search Procedure for Barrier Functions," *Mathematical Programming* (to appear).
6. ———, ——— and Tamir, A., "Nonlinear Programming using Exterior Penalty Functions," report prepared for ONR constract number N0014–67–A–0010, Case Western Reserve University, November 1972.
7. Lootsma, F. A., "Penalty-Function Performance of Several Unconstrained Minimization Techniques," *Phillips Res. Repts.*, Vol. 27 (1972), pp. 358–385.
8. Luenberger, D. G., *Introduction to Linear and Nonlinear Programming*, Addison Wesley, 1973, §§7.2, 7.3.
9. Ortega, J. M. and Rheinboldt, W. C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, 1970.
10. Ostrowski, A., *Solution of Equations and Systems of Equations*, Academic Press, New York, 1966, 2nd ed.
11. Ralston, A., "On Differentiating Error Terms," *Amer. Math. Monthly*, Vol. 70 (1963), pp. 187–189.
12. Traub, J. F., *Iterative Methods for Solution of Equations*, Prentice-Hall, 1964.