JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 27, No. 2, FEBRUARY 1979

Rates of Convergence of a One-Dimensional Search Based on Interpolating Polynomials

A. TAMIR[1]

Communicated by M. Avriel

Abstract. In this study, we derive the order of convergence of line search techniques based on fitting polynomials, using function values as well as information on the smoothness of the function. Specifically, it is shown that, if the interpolating polynomial is based on the values of the function and its first $s-1$ derivatives at $n+1$ approximating points, the rate of convergence is equal to the unique positive root r_{n+1} of the polynomial

$$D_{n+1}(z)=z^{n+1}-(s-1)z^{n}-s\sum_{j=1}^{n} z^{n-j}.$$

For all n, r_n is bounded between s and $s+1$, which in turn implies that the rate can be increased by as much as one wishes, provided sufficient information on the smoothness is incorporated.

Key Words. Mathematical programming, line search procedures, interpolating polynomials, convergence rates.

1. Introduction

Most of the widely used algorithms for solving multidimensional unconstrained minimization problems utilize a one-dimensional search along a direction generated by the algorithm. Computational experience has indicated that a significant portion of the total computational effort is spent in this search.

As reflected in the literature, the most common one-dimensional search techniques used for unconstrained minimization (Refs. 1–4) are based on polynomial interpolation of the objective function. In fact, the methods mentioned above use only low-order polynomials (e.g., quadratic, cubic) for the sequential fitting.

[1] Senior Lecturer of Statistics, Tel Aviv University, Tel Aviv, Israel.

The motivation for our research came from these efficient low-order polynomial interpolations, but the main objective of this study was to develop the general theory of the sequential polynomial fitting as related to the position of proper optima. The emphasis is on convergence and orders of convergence. The special case where the interpolating polynomials do not utilize information on the smoothness of the function to be minimized is reported in Ref. 5.

We start with the description of the sequential polynomial fitting algorithm and then follow with the analysis on properties of convergence and rates of convergence. The paper is concluded with a summary and a short discussion on the practicality of using high-order polynomials for interpolation.

The algorithm studied in the 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 0, and let $x_i, x_{i-1}, \ldots, x_{i-n}$ be $n+1$ distinct approximations to α. If

$$q = \sum_{j=0}^{n} \gamma_j,$$

then there exists a unique polynomial $P_{n,\gamma_0,\gamma_1,\ldots,\gamma_n}$ of degree less than or equal to $q-1$ which satisfies

$$\begin{gathered} P^{(k_j)}_{n,\gamma_0,\ldots,\gamma_n}(x_{i-j}) = f^{(k_j)}(x_{i-j}), \qquad j = 0, 1, \ldots, n, \\ k_j = 0, 1, \ldots, \gamma_j - 1, \qquad \gamma_j \geq 1, \qquad x_{i-k} \neq x_{i-l} \quad \text{if } k \neq l. \end{gathered} \tag{2}$$

For brevity, we write

$$P_{n,\gamma} \equiv P_{n,\gamma_0,\gamma_1,\ldots,\gamma_n}, \qquad P_{n,s} \equiv P_{n,s,s,\ldots,s}$$

where γ signifies the vector $\gamma_0, \gamma_1, \ldots, \gamma_n$. $P_{n,\gamma}$ is called the interpolatory polynomial for f. Then, the new approximation to α, x_{i+1}, is chosen to satisfy

$$P'_{n,\gamma}(x_{i+1}) = 0. \tag{3}$$

If $x_{i+1} = \alpha$, terminate. Otherwise, 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).

The case where only function values are used (i.e., $\gamma_j = 1$, $j = 0, 1, \ldots, n$) is studied in Ref. 5. There, we show that, if the initial $n+1$ approximations are sufficiently close to α, then the sequence generated by the SPFA converges to α. Furthermore, if the sequence is infinite (i.e., convergence is not in a finite number of steps), then the order of convergence is shown to be equal to the unique positive root σ_{n+1} of the polynomial

$$C_{n+1}(z) = z^{n+1} - \sum_{j=1}^{n} z^{n-j}. \tag{4}$$

The sequence $\{\sigma_n\}$ is increasing, approaching the golden section ratio

$$\tau = (1 + \sqrt{5})/2$$

as n approaches infinity. In this work, we extend the above result to the case where information on the smoothness of the function f is included in the interpolating polynomial. Specifically, we will specialize to the equal information case where

$$\gamma_j = s, \qquad j = 0, 1, \dots, n.$$

Since the case $s = 1$ is explored in Ref. 5, we assume throughout this work that $s \geq 2$.

2. Convergence and Convergence Rates

In this work, speed of convergence of line search methods is measured in terms of the following concepts (see Refs. 6 and 7).

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

$$0 \leq \overline{\lim_{k \to \infty}} \, (|e_{k+1}|/|e_k|^p) < \infty.$$

The case 0/0 is regarded as finite. The average order of convergence is the infimum of the numbers $p > 1$ such that

$$\overline{\lim_{k \to \infty}} \, |e_k|^{1/p^k} = 1.$$

The order is infinity if the equality holds for no $p > 1$. Let

$$J = \{x \mid |x - \alpha| \leq L\}. \tag{5}$$

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

Assumption 2.1. (i) If $q = s(n+1)$, where $s \geq 2$ is an integer, then $f^{(q+1)}$ is continuous on J.

(ii) $f^{(2)}(x) \neq 0$ for all $x \in J$. Note that this is equivalent to $f^{(2)}(x) > 0$ for all $x \in J$, since α is an isolated minimum.

(iii) $f^{(q)}(x) \neq 0$ for all $x \in J$.

(iv) If we define constants M_0, M_1, M_2 by

$$M_0 = \min_{x \in J} |f^{(2)}(x)|, \qquad M_1 = \max_{x \in J} |f^{(q)}(x)/q!|,$$

$$M_2 = \max_{x \in J} |f^{(q+1)}(x)/(q+1)!|,$$

then the interval width L in (5) is small enough to satisfy

$$L \leq \tfrac{1}{4}, \tag{6}$$

$$(M_1 q/M_0 L)(L^s + L)^n (L + L^{s-1}) + (M_2/M_0 L)(L^s + L)^{n+1} \leq \tfrac{1}{2}, \tag{7}$$

$$\Gamma = L\{2(M_1 q/M_0 + M_2 L/M_0)\}^{1/(q-2)} \leq 1. \tag{8}$$

The main result of this work is the following theorem.

Theorem 2.1. Under Assumption 2.1, the order of convergence of the SPFA for the equal information case (i.e., $\gamma_j = s \geq 2, j = 0, 1, \ldots, n$) is equal to the unique positive root r_{n+1} of the polynomial

$$D_{n+1}(z) = z^{n+1} - (s-1)z^n - s \sum_{j=1}^{n} z^{n-j}.$$

For all n, $s \leq r_n$, and the sequence of roots $\{r_n\}$ is increasing, approaching $[s + \sqrt{(s^2+4)}]/2$ as n approaches infinity.

In the remainder of this section, we give a number of results leading to a proof of Theorem 2.1. The following two theorems, proved in Appendix A, ensure that the sequence $\{x_i\}$ is well defined and converges to the minimal point α.

Theorem 2.2. Define

$$J = \{x \mid |x - \alpha| \leq L\},$$

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 $P_{n,s}(x)$ of degree $\leq q - 1 = s(n+1) - 1$ satisfying:

$$P_{n,s}^{(k)}(x_{i-j}) = f^{(k)}(x_{i-j}), \qquad j = 0, 1, 2, \ldots, n,$$
$$k = 0, 1, \ldots, s-1, \qquad s \geq 2, \qquad x_{i-t} \neq x_{i-l} \quad \text{if } t \neq l. \tag{9}$$

If f and J satisfy Assumption 2.1, then $P'_{n,s}(x)$ has a real root in J.

Theorem 2.3. Suppose that the conditions of Theorem 2.2 hold and let x_{i+1} in J be a real root of the derivative of the interpolatory polynomial $P_{n,s}(x)$ determined by $x_i, x_{i-1}, \ldots, x_{i-n}$. Then, the sequence $\{x_k\}$ converges to α and

$$|e_k| = |x_k - \alpha| \leq K\Gamma^{r(q,n,k)} \tag{10}$$

for some constant K, $\Gamma < 1$ [defined in (8)], and

$$r(q, n, k) = (q-1)^{k/(n+1)}. \tag{11}$$

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

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

$$P'_{n,s}(x) = f'(x) - [sf^{(q)}(\xi(x))/q!] \sum_{k=0}^{n} (x - x_{i-k})^{s-1} \prod_{\substack{j=0 \\ j\neq k}}^{n} (x - x_{i-j})^s$$
$$-[f^{(q+1)}(\eta(x))/(q+1)!] \prod_{j=0}^{n} (x - x_{i-j})^s, \tag{12}$$

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 (12), and using the relations

$$P'_{n,s}(x_{i+1}) = 0, \qquad (x_{i+1} - x_{i-j}) = (e_{i+1} - e_{i-j}),$$
$$f'(x_{i+1}) = e_{i+1} f^{(2)}(\theta(x_{i+1})),$$

where $\theta(x_{i+1})$ is in the interval $[x_{i+1}, \alpha]$, yields

$$e_{i+1} f^{(2)}(\theta(x_{i+1})) = [sf^{(q)}(\xi(x_{i+1}))/q!] \sum_{k=0}^{n} (e_{i+1} - e_{i-k})^{s-1}$$
$$\times \prod_{\substack{j=0 \\ j\neq k}}^{n} (e_{i+1} - e_{i-j})^s + [f^{(q+1)}(\eta(x_{i+1}))/(q+1)!] \prod_{j=0}^{n} (e_{i+1} - e_{i-j})^s. \tag{13}$$

To find the rate of convergence, we suppose that the SPFA does not terminate in a finite number of steps, i.e., $e_i \neq 0$ for all i, or equivalently none of the approximating points x_i is the sought-for minimum point α.

Recalling that $s \geq 2$, we use (13) to note that

$$|e_{i+1}|/|e_{i+1} - e_i| \xrightarrow[i\to\infty]{} 0, \tag{14}$$

which in turn implies that the order of convergence of the sequence $\{e_i\}$ is at least superlinear. To derive the exact order, we apply (13) to have:

$$\begin{aligned} e_{i+1}f^{(2)}(\theta(x_{i+1})) &= e_i^{s-1} \prod_{j=1}^{n} e_{i-j}^s \Bigg\{ \frac{sf^{(q)}(\xi(x_{i+1}))}{q!} \Bigg[\left(\frac{e_{i+1}}{e_i}-1\right)^{s-1} \prod_{j=1}^{n} \left(\frac{e_{i+1}}{e_{i-j}}-1\right)^s \\ &+ \sum_{k=1}^{n} \left(\frac{e_{i+1}}{e_{i-k}}-\frac{e_i}{e_{i-k}}\right)\left(\frac{e_{i+1}}{e_i}-1\right)^{s-1}\left(\frac{e_{i+1}}{e_{i-k}}-1\right)^{s-1} \prod_{\substack{j=1\\ j\neq k}}^{n} \left(\frac{e_{i+1}}{e_{i-j}}-1\right)^s \Bigg] \\ &+ \frac{f^{(q+1)}(\eta(x_{i+1}))}{(q+1)!}(e_{i+1}-e_i)\left(\frac{e_{i+1}}{e_i}-1\right)^{s-1} \prod_{j=1}^{n} \left(\frac{e_{i+1}}{e_{i-j}}-1\right)^s \Bigg\}. \end{aligned}$$

Use the superlinear convergence of the sequence $\{e_i\}$ and define A_{i+1} by

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

to note that

$$A_{i+1} \to (-1)^{q-1} sf^{(q)}(\alpha)/q!f^{(2)}(\alpha) = A.$$

By Assumption 2.1, $A \neq 0$.

We now use the difference equation (15) to show that the order of convergence of the sequence $\{e_i\}$ is the unique positive real root r_{n+1} of the polynomial

$$D_{n+1}(z) = z^{n+1} - (s-1)z^n - s \sum_{j=1}^{n} z^{n-j}. \tag{16}$$

We need the following lemma (Ref. 8, p. 92).

Lemma 2.1. Consider the linear difference equation

$$u_{i+1} = k_{i+1} + \sum_{j=0}^{n} a_j u_{i-j}, \qquad i = n, n+1, \ldots,$$

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 $t_1, \ldots, t_{n+1}$ be the roots of $Q(x)$, with

$$|t_1| \geq |t_2| \geq \cdots \geq |t_{n+1}|.$$

Assume that

$$|t_1|>1>|t_2|$$

and that, for some u, $0\leq u<|t_1|$,

$$k_i=O(u^i),$$

which means that $|k_i|/u^i\to c$ for some constant c as $i\to\infty$. Then, there exists α_1 such that, as $i\to\infty$,

$$u_i/t_1^i\to\alpha_1.$$

In addition, if $u>|t_2|$,

$$u_i=\alpha_1t_1^i+O(u^i).$$

If $u=|t_2|$ and m is the maximum multiplicity of all zeros of $Q(x)$ with modulus $|t_2|$, then

$$u_i=\alpha_1t_1^i+O(i^m|t_2|^i).$$

A careful examination of the proof in Ref. 8 shows that Lemma 2.1 is true even if the condition

$$|t_1|>1>|t_2|$$

is replaced by the weaker condition

$$|t_1|>1,\qquad |t_1|>|t_2|.$$

Taking absolute values and logarithms of (15), and defining

$$d_i=\log|e_i|\qquad\text{and}\qquad b_i=\log|A_i|,$$

we obtain

$$d_{i+1}=B_{i+1}+s\sum_{j=1}^{n}d_{i-j}+(s-1)d_i,\qquad i=n,n+1,\ldots.$$

Further, defining

$$u_i=d_i/(\log|A|+S),\qquad k_i=B_i/(\log|A|+S),$$

where $S=-1$ if $|A|<1$ and $S=1$ otherwise, yields

$$u_{i+1}=k_{i+1}+s\sum_{j=1}^{n}u_{i-j}+(s-1)u_i,\qquad i=n,n+1,\ldots,\tag{17}$$

where, for i sufficiently large,

$$|k_{i+1}|<1.\tag{18}$$

The characteristic polynomial of (17) is $D_{n+1}(z)$ in (16). Consider first the case where $n+1$ is odd. It is shown in Appendix B that, in this case, the roots of $D_{n+1}(z)$ satisfy

$$|t_1|>1>|t_2|.$$

By (18), we can apply Lemma 2.1 with $u=1$ to obtain

$$u_i=\alpha_1 t_1^i+O(1),$$

$$|e_i|=\exp\{-\beta_1 t_1^i+O(1)\},$$

where $\beta_1>0$, since $|e_i|\to 0$. This implies that

$$|e_{i+1}|/|e_i|^t=\exp\{\beta_1 t_1^i(t-t_1)+O_1(1)+tO_2(1)\},$$

which yields that the order of convergence of the sequence $\{e_i\}$ is t_1. Also note that the average order of convergence is t_1. Suppose now that $n+1$ is even. Then, from Appendix B,

$$t_1>1 \qquad \text{and} \qquad t_2=-1.$$

The comment following Lemma 2.1 justifies its use in this circumstance; and, using $u=|t_2|=1$, we obtain

$$u_i=\alpha_1 t_1^i+O(i),$$

which implies that

$$|e_i|=\exp\{\gamma_1 t_1^i+O(i)\}. \tag{19}$$

Since $|e_i|\to 0$,

$$\gamma_1\leq 0.$$

If $\gamma_1=0$, then

$$|e_i|=\exp\{O(i)\},$$

which contradicts (10). Hence,

$$\gamma_1<0.$$

It is then easily verified that (19) implies that the order of convergence of the sequence $\{e_i\}$ as well as the average order are again t_1. Theorem 2.1 follows from the preceding discussion and Appendix B.

3. Summary and Comments

We have shown that, if the interpolating polynomial is based on the values of the function and its first $s-1$ derivatives at $n+1$ approximating points, then the rate of convergence is the unique positive root $r_{n+1}(s)$ of the polynomial $D_{n+1}(z)$ (see Theorem 2.1). Furthermore, $r_n(s)$ is bounded between s and $s+1$, which in turn implies that the rate can be increased by as much as one wishes, provided sufficient information on the smoothness is incorporated.

At this point, we should emphasize that our scheme is not proposed as a computationally practical tool unless the polynomial is of low order. Computational experience has shown that the increase in the theoretical rate of convergence does not compensate for the extra computational effort involved in dealing with high-order polynomials. In this respect, it is worthwhile to mention several efficiency measures used in the literature, which also supports the above computational observation. Since our procedure requires s new pieces of information per iteration, the information efficiency is equal to $r_n(s)/s$, while Ostrowski's efficiency index becomes $(r_n(s))^{1/s}$ (see Ref. 9, p. 11–12). It is then easily verified that both measures of efficiency decrease to 1 as s increases. Note, however, that the above measures presume the same *cost* for evaluating any piece of information, e.g., evaluating $f(x)$ has the same effect as calculating $f^{(1)}(x)$.

Finally, we point out the main difference between our method and the one obtained from a direct interpolation of $f^{(1)}(x)$. It is well known (Ref. 9, Section 3.3) that, in the latter case, the orders of convergence tend to $1+s$, which is larger than $[s+\sqrt{(s^2+4)}]/2$, the limit achieved by the scheme presented in this paper. But observe that, although both techniques use s new pieces of information per iteration, these pieces are not identical. While the direct approach involves the calculation of $f^{(1)}, f^{(2)}, \ldots, f^{(s)}$ at a certain point, our scheme evaluates the function itself $f^{(0)}$ and $f^{(1)}$, $f^{(2)}, \ldots, f^{(s-1)}$. In particular, the SPFA can be applied even when the expression for $f^{(1)}(x)$ is not given analytically ($s=1$). Thus, the application of the direct approach requires the calculation of a derivative of higher dimension.

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

In this appendix, we prove Theorems 2.2 and 2.3 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 2.2. Since $f^{(q)}(x)$ is continuous, it is well known (e.g., Ref. 9, p. 61) that

$$f(x) = P_{n,s}(x) + [f^{(q)}(\xi(x))/q!] \prod_{j=0}^{n} (x - x_{i-j})^s \tag{20}$$

where $\xi(x)$ lies in the interval determined by $x_i, x_{i-1}, \ldots, x_{i-n}, x$. To derive an expression for $P'_{n,s}(x)$, we apply a result due to Ralston (Ref. 10), which states that

$$(1/q!)(d/dx)f^{(q)}(\xi(x)) = [1/(q+1)!]f^{(q+1)}(\eta(x)), \tag{21}$$

where $\eta(x)$ is again a mean value in the interval of interpolation. Differentiating (20) and using (21) yields

$$P'_{n,s}(x) = f'(x) - [sf^{(q)}(\xi(x))/q!] \sum_{k=0}^{n} (x - x_{i-k})^{s-1} \prod_{\substack{j=0 \\ j\neq k}}^{n} (x - x_{i-j})^s$$
$$- [f^{(q+1)}(\eta(x))/(q+1)!] \prod_{j=0}^{n} (x - x_{i-j})^s. \tag{22}$$

We now show that, under the assumptions of the theorem, $P'_{n,s}(x)$ has a zero in J. Note first that $f^{(2)}(x) > 0$, $\forall x \in J$, since α is a minimum point, and hence $f^{(2)}(\alpha) \geq 0$. The theorem follows when we prove that

$$P'_{n,s}(\alpha - L) < 0 \qquad \text{and} \qquad P'_{n,s}(\alpha + L) > 0.$$

$f'(\alpha) = 0$ implies that

$$f'(x) = f'(x) - f'(\alpha) = (x - \alpha) f^{(2)}(\gamma(x)),$$

where $\gamma(x)$ is in J. Substituting $x = \alpha - L$ in (22) yields

$$P'_{n,s}(\alpha - L) = -Lf^{(2)}(\gamma(\alpha - L)) - [sf^{(q)}(\xi(\alpha - L))/q!]$$
$$\times \sum_{k=0}^{n} (\alpha - L - x_{i-k})^{s-1} \prod_{\substack{j=0 \\ j\neq k}}^{n} (\alpha - L - x_{i-j})^s$$
$$- [f^{(q+1)}(\eta(\alpha - L))/(q+1)!] \prod_{j=0}^{n} (\alpha - L - x_{i-j})^s.$$

$P'_{n,s}(\alpha - L)$ is negative if

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

To prove that $T < 1$, we note that

$$T \leq |T| \leq (M_2/M_0L)(2L)^{s(n+1)} + (M_1q/M_0L)(2L)^{s(n+1)-1}.$$

Using (6), we observe that $(2L)^k \leq L$ for all $k \geq 2$, which in turn yields the following inequalities

$$(2L)^{s(n+1)} \leq L^{(n+1)} \leq (L+L^s)^{n+1},$$
$$(2L)^{sn} \leq (L+L^s)^n, \qquad (L+L)^{s-1} \leq L+L^{s-1}.$$

Thus,

$$\begin{aligned} T &\leq (M_2/M_0L)(2L)^{s(n+1)} + (M_1q/M_0L)(2L)^{s(n+1)-1} \\ &\leq (M_2/M_0L)(L+L^s)^{n+1} + (M_1q/M_0L)(L+L^s)^n(L+L^{s-1}). \end{aligned}$$

We finally apply our assumption (7) to obtain $T<1$.

Similar arguments lead to the conclusion that $P'_{n,s}(\alpha+L)>0$, and hence the theorem follows.

Proof of Theorem 2.3. Substituting $x = x_{i+1}$ in (22), we obtain

$$\begin{aligned} f'(x_{i+1}) = [sf^{(q)}(\theta_1)/q!] \sum_{l=0}^{n} (x_{i+1}-x_{i-l})^{s-1} \prod_{\substack{j=0 \\ j\neq l}}^{n} (x_{i+1}-x_{i-j})^s \\ + [f^{(q+1)}(\theta_2)/(q+1)!] \prod_{j=0}^{n} (x_{i+1}-x_{i-j})^s, \end{aligned}$$

where

$$\theta_1 = \xi(x_{i+1}), \qquad \theta_2 = \eta(x_{i+1}).$$

Defining

$$e_k = x_k - \alpha, \qquad k = 1, 2, \ldots,$$

and noting that

$$f'(x_{i+1}) = e_{i+1} f^{(2)}(\theta_3), \qquad \theta_3 = \gamma(x_{i+1}),$$

yields

$$\begin{aligned} M_0|e_{i+1}| \leq sM_1 \sum_{l=0}^{n} |e_{i+1}-e_{i-l}|^{s-1} \prod_{j\neq l} |e_{i+1}-e_{i-j}|^s \\ + M_2 \prod_{j=0}^{n} |e_{i+1}-e_{i-j}|^s. \end{aligned} \tag{23}$$

Let $m \geq 1$ be integer. Then,

$$|e_{i+1}-e_{i-j}|^m \leq |e_{i+1}|L^{m-1}(2^m-1) + |e_{i-j}|^m \leq |e_{i+1}| + |e_{i-j}|^m, \tag{24}$$

where the right inequality is implied by the assumption (6). Applying (24) to (23) results in

$$\begin{aligned} M_0|e_{i+1}| &\leq sM_1(n+1)(|e_{i+1}| + \max_{0\leq j\leq n} |e_{i-j}|^s)^n(|e_{i+1}| + \max_{0\leq j\leq n} |e_{i-j}|^{s-1}) \\ &\quad + M_2(|e_{i+1}| + \max_{0\leq j\leq n} |e_{i-j}|^s)^{n+1} \\ &\leq sM_1(n+1)\{|e_{i+1}|(L^{sn} + (1/L)[(L^s+L)^n - L^{sn}](L+L^{s-1})) \\ &\quad + \max_{0\leq j\leq n} |e_{i-j}|^{s(n+1)-1}\} \\ &\quad + M_2\{(|e_{i+1}|/L)[(L^s+L)^{n+1} - L^{s(n+1)}] + \max_{0\leq j\leq n} |e_{i-j}|^{s(n+1)}\}. \end{aligned}$$

Hence,

$$\begin{aligned} |e_{i+1}| &\leq \{[sM_1(n+1)/M_0L](L^s+L)^n(L+L^{s-1}) + (M_2/LM_0)(L^s+L)^{n+1}\}|e_{i+1}| \\ &\quad + \{sM_1(n+1)/M_0 + M_2L/M_0\} \max_{0\leq j\leq n} |e_{i-j}|^{q-1}. \qquad (25) \end{aligned}$$

By Assumption 2.1, (7), we have

$$|e_{i+1}| \leq C \max_{0\leq j\leq n} |e_{i-j}|^{q-1},$$

where

$$C = 2(M_1q/M_0 + M_2L/M_0).$$

For any positive integer k, define

$$t_k = |e_k|C^{1/(q-2)}.$$

Then, (25) yields

$$t_{i+1} \leq \max_{0\leq j\leq n} t_{i-j}^{q-1}.$$

Let

$$\Gamma = LC^{1/(q-2)}.$$

It can be verified by induction that, if

$$k = t(n+1) + l, \qquad t \geq 1, \qquad l = 0, 1, \ldots, n,$$

then

$$t_k \leq \Gamma^{(q-1)^t}.$$

Letting

$$r(q, n, k) = (q-1)^{k/(n+1)}$$

and observing that $\Gamma < 1$ and

$$t = k/(n+1) - l/(n+1)$$

yields

$$|e_k| = t_k C^{-1/(q-2)} \leq C^{-1/(q-2)} \Gamma^{r(q,n,k)}.$$

5. Appendix B: Roots of the Indicial Equation

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

$$D_k(z) = z^k - (s-1)z^{k-1} - s(z^{k-2} + z^{k-3} + \cdots 1), \tag{26}$$

when $k \geq 2$ and $s \geq 2$. We show that $D_k(z)$ has a unique simple positive root r_k, with modulus greater than one, and that all other roots are also simple with moduli less than or equal to one. In fact, it will be proved that, if k is odd, r_k is the only real root and that the other $k-1$ roots are inside the unit disc. When k is even, $z = -1$ and r_k are the only real roots and the other $k-2$ roots have moduli less than one. It is finally demonstrated that the sequence $\{r_k\}$, $k = 2, 3, \ldots$, is increasing and tends to $[s + \sqrt{(s^2+4)}]/2$.

Lemma 5.1. Let $D_k(z)$, $k \geq 3$, be defined by (26). $D_k(z)$ has a unique simple positive root r_k,

$$s < r_k < [s + \sqrt{(s^2+4)}]/2.$$

If k is odd, r_k is the only real root; and, if k is even, $z = -1$ is the only other real root of $D_k(z)$ and is simple.

Proof. Observe that

$$D_k(z) = [1/(z-1)][z^{k-1}(z^2 - sz - 1) + s]. \tag{27}$$

We verify that $s \geq 1$ implies that $D_k(s) < 0$. Furthermore, $D_k(z)$ is positive at $(s + \sqrt{(s^2+4)}/2$, and thus there exists r_k,

$$s < r_k < [s + \sqrt{(s^2+4)}]/2,$$

and $D_k(r_k) = 0$. To see that r_k is simple and also the unique positive root, note first that

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

$$a_k = s/r_k,$$

$$a_i = (1/r_k)(a_{i+1} + s), \qquad 2 \leq i < k.$$

Thus, $a_i > 0$, $i = 2, 3, \ldots, k$, and the result follows. Let k be even. Then,

$$D_k(z) = (z+1)(z^{k-1} - sz^{k-2} - sz^{k-4} \cdots - s).$$

Hence, $z = -1$ is a simple root. We then verify that $D_k(t) < 0$ for $-1 < t \leq 0$ and $D_k(t) > 0$ for $t < -1$, to obtain that $z = -1$ is the only nonpositive root. Suppose now that $k \geq 3$ is odd. We have

$$D_k(z) = z^k - (s-1)z^{k-1} - s(z+1)(z^{k-3} + z^{k-5} + \cdots + 1). \qquad (28)$$

Therefore, $D_k(t) < 0$ for $-1 \leq t \leq 0$. Noting that $t^2 - st - 1$ is positive for $t < -1$, we use (27) to have that $D_k(t) < 0$ for $t < -1$. Thus, $D_k(t) < 0$ for all $t \leq 0$, and r_k is the unique real root.

Lemma 5.2. All the roots of $D_k(z)$ are simple.

Proof. If $z \neq 1$ is a multiple root of $D_k(z)$, it is also a multiple root of

$$E_K(z) = (z-1)D_k(z).$$

A multiple root of $E_K(z)$ is a root of $E'_K(z)$, which implies from (27) that

$$(k+1)z^k - skz^{k-1} - (k-1)z^{k-2} = 0.$$

$z = 0$ is not a root, and we have

$$(k+1)z^2 - skz - (k-1) = 0,$$

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 $\{r_k\}$ is increasing.

Lemma 5.3. $\{r_k\}$, $k = 2, 3, \ldots$, is an increasing sequence, and

$$\lim_k r_k = [s + (\sqrt{s^2 + 4})]/2.$$

Proof. Using Lemma 5.1, the monotonicity will follow if we show that

$$D_k(r_{k-1}) < 0.$$

From (27), we get

$$(z-1)D_k(z) - s = z[(z-1)D_{k-1}(z) - s].$$

Hence,

$$(r_{k-1} - 1)D_k(r_{k-1}) - s = -r_{k-1}s \qquad \text{and} \qquad D_k(r_{k-1}) = -s.$$

The sequence $\{r_k\}$ is a bounded increasing sequence, and hence

$$\lim_k r_k = \beta$$

exists.

$$r_k^{k-1}(r_k^2 - sr_k - 1) = -s, \qquad 1 < r_k$$

$$\Rightarrow \beta^2 - s\beta - 1 = 0 \qquad \text{and} \qquad \beta = [s + \sqrt{(s^2+4)}]/2.$$

To prove that the no real roots of $D_k(z)$ have moduli less than or equal to 1, we need the following two results.

Theorem 5.1. (*Ref. 9, p. 51*). Let

$$f_k(z) = z^k - a(z^{k-1} + z^{k-2} + \cdots + 1), \qquad ka > 1, \qquad \text{and} \qquad k \geq 2.$$

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 5.4. (*Ref. 8, 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 zeros 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 5.2. If k is odd, the $k-1$ roots of $D_k(z)/(z - r_k)$ have moduli <1. If k is even, the $k-2$ roots of $D_k(z)/(z - r_k)(z+1)$ have moduli <1.

Proof. Let $\mathscr{E} > 0$ be arbitrarily small, and consider the polynomial $D_k(z) - tz^{k-1}$ for $t \in [\mathscr{E}, 1]$. We show that

$$D_k(z) - tz^{k-1} \neq 0 \qquad \text{for all } z \in \{z \,|\, |z| = 1\}.$$

Since

$$D_k(1) - t < 0 \qquad \text{for } \mathscr{E} \leq t \leq 1,$$

it is sufficient to prove that

$$(z-1)\{D_k(z) - tz^{k-1}\} \neq 0 \qquad \text{for all } z \neq 1 \text{ and } |z| = 1.$$

Suppose that

$$(z-1)\{D_k(z) - tz^{k-1}\} = 0$$

for some $z \neq 1$ and $|z| = 1$. Then,

$$z^{k-1}[z^2 - z(s+t) - (1-t)] + s = 0$$
$$\Rightarrow |z^2 - z(s+t) - (1-t)| = s.$$

If

$$z = \cos\theta + i\sin\theta,$$

then

$$[\cos 2\theta - (s+t)\cos\theta - (1-t)]^2 + [\sin 2\theta - (s+t)\sin\theta]^2 = s^2,$$

which yields

$$-2(1-t)\cos^2\theta - t(s+t)\cos\theta + t^2 + t(s-2) + 2 = 0.$$

Let $y = \cos\theta$. Then, $y = 1$ is one root of the quadratic

$$2(1-t)y^2 + t(s+t)y - (t^2 + t(s-2) + 2) = 0. \tag{29}$$

For $t = 1$, $y = 1$ is the only root, and we obtain $\cos\theta = 1$, which contradicts $z \neq 1$. Let $t \in [\mathscr{E}, 1)$. Then, the second root of (29) is

$$y(t) = -[t^2 + t(s-2) + 2]/2(1-t) = (-t^2 - ts)/2(1-t) - 1 < -1.$$

Thus, we have the contradiction $\cos\theta < -1$, and we get that

$$D_k(z) - tz^{k-1} \neq 0 \qquad \text{for } z \in \{z \,|\, |z| = 1\}.$$

Observing that, for $t = 1$, $D_k(z) - tz^{k-1}$ yields the polynomial $f_k(z)$ with $a = s$ discussed in Theorem 5.1, we apply Lemma 5.4 to conclude that, for any positive t arbitrarily close to zero, the polynomial $D_k(z) - tz^{k-1}$ has $k-1$ roots *inside* the disc $\{z \,|\, |z| \leq 1\}$. Continuity arguments (see, for example, Ref. 8, Appendix A) lead to the conclusion that $D_k(z)$ has $k-1$ roots in $\{z \,|\, |z| \leq 1\}$. By substituting $t = 0$ in (29), we easily verify that the only possible root of $D_k(z)$ on the boundary of the disc is $z = -1$, which is a root iff k is even. Hence, the theorem is proved.

References

1. FLETCHER, R., and POWELL, M. J. D., *A Rapidly Convergent Descent Method for Minimization*, Computer Journal, Vol. 6, pp. 163–168, 1963.
2. FLETCHER, R., and REEVES, C., *Function Minimization by Conjugate Gradients*, Computer Journal, Vol. 7, pp. 149–154, 1964.
3. GOLDFARB, D., *Extension of Davidon's Variable Metric Method to Maximization under Linear Inequality and Equality Constraints*, SIAM Journal on Applied Mathematics, Vol. 17, pp. 739–764, 1969.

4. LOOTSMA, P. A., *Penalty-Function Performance of Several Unconstrained Minimization Techniques*, Phillips Research Report, Vol. 27, pp. 358–385, 1972.
5. TAMIR, A., *A One-Dimensional Search Based on Interpolating Polynomials Using Function Values Only*, Management Science (to appear).
6. LUENBERGER, D. G., *Introduction to Linear and Nonlinear Programming*, Sections 7.2 and 7.3, Addison-Wesley Publishing Company, Reading, Massachusetts, 1973.
7. ORTEGA, J. M., and RHEINBOLDT, W. C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, New York, 1970.
8. OSTROWSKI, A., *Solution of Equations and Systems of Equations*, Academic Press, New York, New York, 1966.
9. TRAUB, J. F., *Iterative Methods for Solution of Equations*, Prentice-Hall, Englewood Cliffs, New Jersey, 1964.
10. RALSTON, A., *On Differentiating Error Terms*, American Mathematical Monthly, Vol. 70, pp. 187–189, 1963.