DISCUSSION PAPER NO. 143

RATES OF CONVERGENCE OF A ONE-DIMENSIONAL SEARCH BASED ON INTERPOLATING POLYNOMIALS

by

Arie Tamir

May 1975

Northwestern University
Graduate School of Management
Managerial Economics and Decision Sciences Department

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.

RATES OF CONVERGENCE OF A ONE-DIMENSIONAL SEARCH BASED ON INTERPOLATING POLYNOMIALS

By Arie Tamir

This study is concerned with a line search technique based on interpolating polynomials where the focus is on convergence properties and rates of convergence.

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

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

For brevity we write $P_{n,\gamma} \equiv P_{n,\gamma_0,\gamma_1,\ldots,\gamma_n}$ $P_{n,s} \equiv P_{n,s,s,\ldots,s}$ 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 [5]. There we show that if the initial n+1 approximations are sufficently 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 $\tau = \frac{1+\sqrt{5}}{2}$ 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 $\gamma_j = s$, $j = 0,1,\ldots,n$. Since the case s = 1 is explored in [5], we assume throughout this work that $s \geq 2$.

Convergence and Convergence Rates

In this work speed of convergence of line search methods is measured in terms of the following concepts. (See [1], [2]).

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

$$0 \le \overline{\lim_{k\to\infty}} \frac{|e_{k+1}|}{|e_k|^p} < \infty .$$

(The case o/o is regarded as finite). The average order of convergence is the infinum 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

$$J = \{x| \; |x-\alpha| \le L\} \tag{5}$$

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

Assumption 1

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

2. $f^{(2)}(x) \neq 0$ 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).

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

4. If we define constants M_0, M_1, M_2 by

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

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

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

$$L \leq \frac{1}{4} \tag{6}$$

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

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

The main result of this work is:

Theorem 1:

Under Assumption 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 $\frac{s + \sqrt{s^2 + 4}}{2}$ as n approaches infinity.

In the remainder of this section, we give a number of results leading to a proof of Theorem 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.

Theorem 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:

$$P^{(k)}_{n,s}(x_{i-j}) = f^{(k)}(x_{i-j}) \qquad j = 0,1,2,\ldots,n. \tag{9}$$

$$k = 0,1,\ldots,s-1, \quad s \geq 2; \quad x_{i-t} \neq x_{i-\ell} \text{ if } t \neq \ell$$

If f and J satisfy Assumption 1 then $P'_{n,s}(x)$ 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 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

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

for some constant K. $\Gamma < 1$ (defined in (8)) and

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

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

$$P'_{n,s}(x) = f'(x) - \frac{sf^{(q)}(\xi(x))}{q!}\sum_{k=0}^{n}(x-x_{i-k})^{s-1}\prod_{\substack{j=0\\ j\neq k}}^{n}(x-x_{i-j})^{s}$$

$$- \frac{f^{(q+1)}(\eta(x))}{(q+1)!}\prod_{j=0}^{n}(x-x_{i-j})^{s} \qquad (12)$$

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

$$P'_{n,s}(x_{i+1}) = 0, \quad (x_{i+1}-x_{i-j}) = (e_{i+1}-e_{i-j})$$

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

$$e_{i+1}f^{(2)}(\theta(x_{i+1})) = \frac{sf^{(q)}(\xi(x_{i+1}))}{q!}\sum_{k=0}^{n}(e_{i+1}-e_{i-k})^{s-1} \prod_{\substack{j=0\\ j\neq k}}^{n}(e_{i+1}-e_{i-j})^{s} + \frac{f^{(q+1)}(\eta(x_{i+1}))}{(q+1)!}\prod_{j=0}^{n}(e_{i+1}-e_{i-j})^{s} \qquad (12a)$$

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 (12a) to note that

$$\frac{|e_{i+1}|}{|e_{i+1}-e_i|} \xrightarrow[i\to\infty]{} 0 \qquad (13)$$

which in turn implies that the order of convergence of the sequene $\{e_i\}$ is at least superlinear. To derive the exact order we apply (12a) to have:

$$e_{i+1}\, f^{(2)}(\theta(x_{i+1})) = e_i^{s-1} \prod_{j=1}^{n} e_{i-j}^{s} \Big\{ \frac{s\, f^{(q)}(\xi(x_{i+1}))}{q!} \Big[\Big(\frac{e_{i+1}}{e_i} - 1\Big)^{s-1} \prod_{j=1}^{n} \Big(\frac{e_{i+1}}{e_{i-j}} - 1\Big)^{s} + \sum_{k=1}^{n} \Big(\frac{e_{i+1}}{e_{i-k}} - \frac{e_i}{e_{i-k}}\Big) \Big(\frac{e_{i+1}}{e_i} - 1\Big)^{s-1} \Big(\frac{e_{i+1}}{e_{i-k}} - 1\Big)^{s-1} \prod_{\substack{j=1 \\ j \neq k}}^{n} \Big(\frac{e_{i+1}}{e_{i-j}} - 1\Big)^{s} \Big] +$$

$$+ \frac{f^{(q+1)}(\eta(x_{i+1}))}{(q+1)!} (e_{i+1} - e_i) \cdot \Big(\frac{e_{i+1}}{e_i} - 1\Big)^{s-1} \prod_{j=1}^{n} \Big(\frac{e_{i+1}}{e_{i-j}} - 1\Big)^{s} \Big\} .$$

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

$$e_{i+1} = A_{i+1}\, e_i^{s-1} \prod_{j=1}^{n} e_{i-j}^{s} \tag{14}$$

to note that

$$A_{i+1} \to \frac{(-1)^{q-1} s\, f^{(q)}(\alpha)}{q!\, f^{(2)}(\alpha)} = A$$

By Assumption 1, $A \neq 0$.

We now use the difference equation (14) to show that the order of convergence of the sequence $\{e_i\}$ is the unique positive real root, r_{n+1}, of the polynomial

$$D_{n+1}(z) = z^{n+1} - (s-1)z^{n} - s \sum_{j=1}^{n} z^{n-j} . \tag{15}$$

We need the following lemma [3, p. 92].

Lemma 2:

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, for some u, $0 < u < |t_1|$

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

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

$$\frac{u_i}{t_1^{\,i}} \to \alpha_1.$$

In addition if $u > |t_2|$

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

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

$$u_i = \alpha_1 t_1^i + O(i^m |t_2|^i).$$

A careful examination of the proof in [3] shows that Lemma 2 is true even if the condition $|t_1| > 1 > |t_2|$ is replaced by the weaker condition

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

Taking absolute values and logs of (14), and defining

$$d_i = \ell n \; |e_i| \text{and } B_i = \ell n \; |A_i|$$

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

$$u_i = \frac{d_i}{\ell n|A|+S}, \quad k_i = \frac{B_i}{\ell n|A|+S}$$

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

$$u_{i+1} = k_{i+1} + s\sum_{j=1}^{n} u_{i-j} + (s-1)u_i \qquad i=n, n+1, \ldots \quad (16)$$

where, for i sufficiently large

$$|k_{i+1}| < 1. \quad (17)$$

The characteristic polynomial of (16) is $D_{n+1}(z)$ in (15). 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 (17), we can apply Lemma 2 with $u = 1$ to obtain:

$$u_i = \alpha_1 t_1^i + 0\ (1)$$

and

$$|e_i| = \exp\{-\beta_1 t_1^i + 0(1)\}$$

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

$$\frac{|e_{i+1}|}{|e_i|^t} = \exp\{\beta_1 t_1^i\ (t-t_1) + 0_1(1) + t0_2(1)\}$$

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, $t_1>1$ and $t_2=-1$. The comment following Lemma 2 justifies its use in this circumstance and using $u = |t_2| = 1$ we obtain

$$u_i = \alpha_1 t_1^i + 0\ (i)$$

which implies

$$|e_i| = \exp\{\gamma_1 t_1^i + 0\ (i)\}. \tag{18}$$

Since $|e_i| \to 0, \gamma_1 \leq 0$. If $\gamma_1 = 0$ then $|e_i| = \exp\{0(i)\}$, which contradicts (10). Hence $\gamma_1 < 0$. It is then easily verified that (18) implies that the order of convergence of the sequence $\{e_i\}$ as well as the average order are again t_1. Theorem 1 follows from the preceeding discussion and Appendix B.

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 $f^{(q)}(x)$ is continuous it is well known (e.g. [6 ,P.61]) that

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

where $\xi(x)$ lies in the interval determined by $x_i, x_{i-1}, \ldots, x_{i-n}, x$. To derive an expression for $P'_{n,s}(x)$ we apply a result due to Ralston [4], which states that

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

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

$$P'_{n,s}(x) = f'(x) - \frac{s\, f^{(q)}(\xi(x))}{q!} \sum_{k=0}^{n} (x-x_{i-k})^{s-1} \prod_{\substack{j=0 \\ j\neq k}}^{n} (x-x_{i-j})^{s} -$$

$$\frac{f^{(q+1)}(\eta(x))}{(q+1)!} \prod_{j=0}^{n} (x-x_{i-j})^{s} \tag{A.3}$$

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

$$P'_{n,s}(\alpha-L) = -L\; f^{(2)}(\gamma(\alpha-L)) - \frac{s\; f^{(q)}(\xi(\alpha-L))}{q!} \sum_{k=0}^{n} (\alpha-L-x_{i-k})^{s-1} \prod_{\substack{j=0\\ j\neq k}}^{n} (\alpha-L-x_{i-j})^{s}$$

$$- \frac{f^{(q+1)}(\eta(\alpha-L))}{(q+1)!} \prod_{j=0}^{n} (\alpha-L-x_{i-j})^{s}$$

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

$$T = \frac{1}{Lf^{(2)}(\gamma(\alpha-L))} \Big[-\frac{s\; f^{(q)}(\xi(\alpha-L))}{(q)!} \sum_{k=0}^{n} (\alpha-L-x_{i-k})^{s-1} \prod_{j\neq k} (\alpha-L-x_{i-j})^{s}$$

$$- \frac{f^{(q+1)}(\eta(\alpha-L))}{(q+1)!} \prod_{j=0}^{n} (\alpha-L-x_{i-j})^{s} \Big] < 1.$$

To prove that $T < 1$ we note that $T \leq |T| \leq \frac{M_2}{M_0 L} (2L)^{s(n+1)} + \frac{M_1 q}{M_0 L} (2L)^{s(n+1)-1}$. Using (6) we observe that $(2L)^{k} \leq L$ for all $k \geq 2$, which in turn yields the following inequalities

$$(2L)^{s(n+1)} \leq L^{(n+1)} \leq (L+L^{s})^{n+1}$$

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

Thus

$$T \leq \frac{M_2}{M_0 L}(2L)^{s(n+1)} + \frac{M_1 q}{M_0 L}(2L)^{s(n+1)-1} \leq \frac{M_2}{M_0 L}(L+L^s)^{n+1} +$$

$$\frac{M_1 q}{M_0 L}(L+L^s)^n (L+L^{s-1}).$$

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

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

Proof of Theorem 3. Substituting $x = x_{i+1}$ in (A.3) we obtain

$$f'(x_{i+1}) = \frac{s\, f^{(q)}(\theta_1)}{q!} \sum_{\ell=0}^{n} (x_{i+1}-x_{i-\ell})^{s-1} \prod_{\substack{j=0 \\ j\neq\ell}}^{n} (x_{i+1}-x_{i-j})^s +$$

$$+ \frac{f^{(q+1)}(\theta_2)}{(q+1)!} \prod_{j=0}^{n} (x_{i+1}-x_{i-j})^s$$

where

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

Defining $e_k = x_k - \alpha$, k=1,2,... and noting that

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

yield

$$M_o|e_{i+1}| \leq sM_1 \sum_{\ell=0}^{n} |e_{i+1}-e_{i-\ell}|^{s-1} \prod_{j\neq\ell} |e_{i+1}-e_{i-j}|^s + M_2 \prod_{j=0}^{n} |e_{i+1}-e_{i-j}|^s \quad \text{(A.}$$

Let m ≥ 1 be integer, then

$$|e_{i+1}-e_{i-j}|^m \leq |e_{i+1}|\, L^{m-1}(2^m-1) + |e_{i-j}|^m \leq |e_{i+1}| + |e_{i-j}|^m \qquad \text{(A.}$$

where the right inequality is implied by the assumption (6).

Applying (A.5) to (A.4) results in

$$M_0|e_{i+1}| \le sM_1(n+1)(|e_{i+1}| + \max_{0\le j\le n}|e_{i-j}|^s)^n(|e_{i+1}| + \max_{0\le j\le n}|e_{i-j}|^{s-1}) +$$

$$+ M_2\,(|e_{i+1}| + \max_{0\le j\le n}|e_{i-j}|^s)^{n+1}$$

$$\le sM_1(n+1)\{|e_{i+1}|(L^{sn} + \frac{1}{L}[(L^s+L)^n - L^{sn}](L+L^{s-1}) + \max_{0\le j\le n}|e_{i-j}|^{s(n+1)-1}\}$$

$$+ M_2\,\{\frac{|e_{i+1}|}{L}[(L^s+L)^{n+1} - L^{s(n+1)}] + \max_{0\le j\le n}|e_{i-j}|^{s(n+1)}.$$

Hence,

$$|e_{i+1}| \le \{\frac{sM_1(n+1)}{M_0}\frac{1}{L}(L^s+L)^n(L+L^{s-1}) + \frac{M_2}{LM_0}(L^s+L)^{n+1}\}\;|e_{i+1}|$$

$$+ \{\frac{sM_1(n+1)}{M_0} + \frac{M_2}{M_0}L\}\max_{0\le j\le n}|e_{i-j}|^{q-1} \qquad (A.6)$$

By Assumption 1 (7)

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

where $C = 2(\frac{M_1 q}{M_0} + \frac{M_2}{M_0}L)$.

For any positive integer k define $t_k = |e_k|\,C^{1/(q-2)}$. Then (A.6) yields

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

Let $\Gamma = LC^{1/(q-2)}$, then it can be verified by induction that if $k = t(n+1) + \ell$, $t \ge 1$, $\ell = 0,1,\ldots,n$, then

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

Letting $r(q,n,k) = (q-1)^{k/(n+1)}$ and observing that $\Gamma < 1$ and $t = \frac{k}{n+1} - \frac{\ell}{n+1}$ yield

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

Appendix B

The Roots of the Indicial Equation

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

$$D_k(z) = z^k - (s-1)z^{k-1} - s(z^{k-2} + z^{k-3} +\ldots 1) \qquad \text{(B.1)}$$

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 $\frac{s + \sqrt{s^2+4}}{2}$.

Lemma 1:

Let $D_k(z)$, $k \geq 3$ be defined by (B.1). $D_k(z)$ has a unique simple positive root r_k, $s < r_k < \frac{s + \sqrt{s^2+4}}{2}$. 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:

$$D_k(z) = \frac{1}{z-1}\,[z^{k-1}(z^2-sz-1)\pm s] \qquad \text{(B.2)}$$

We verify that $s \geq 1$ implies $D_k(s) < 0$. Furthermore, $D_k(z)$ is positive at $\frac{s + \sqrt{s^2+4}}{2}$, and thus there exists r_k, $s < r_k < \frac{s + \sqrt{s^2+4}}{2}$ and $D_k(r_k) = 0$. To see that r_k is simple and also the unique positive root note first that

The following lemma shows that the sequence $\{r_k\}$ is increasing.

Lemma 3:

$\{r_k\}$, k=2,3,... is an increasing sequence and $\lim_k r_k = \frac{s+\sqrt{s^2+4}}{2}$.

Proof: Using Lemma 1 the monotonicity will follow if we show that $D_k(r_{k-1}) < 0$. From (B.2) we get

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

Hence $(r_{k-1}-1)\, D_k(r_{k-1})-s = -r_{k-1}\cdot s$ and $D_k(r_{k-1}) = -s$. The sequence $\{r_k\}$ is a bounded increasing sequence and hence $\lim_k r_k = \beta$ exists.

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

$$\Rightarrow \qquad \beta^2 - s\beta - 1 = 0 \quad \text{and} \quad \beta = \frac{s+\sqrt{s^2+4}}{2} .$$

To prove that the none real roots of $D_k(z)$ have moduli less than or equal to 1 we need the following two results.

Theorem 1: (Traub, [6, p.51])

Let $f_k(z) = z^k - a(z^{k-1} + z^{k-2} + \dots + 1)$, $ka > 1$ 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 4: (Ostrowski, [3, 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 inverval $a \leq t \leq b$, 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 2: If k is odd the k-1 roots of $\frac{D_k(z)}{z-r_k}$ have moduli < 1. If k is even the k-2 roots of $\frac{D_k(z)}{(z-r_k)(z+1)}$ have moduli < 1.

Proof: Let $\varepsilon > 0$ be arbitrary small and consider the polynomial $D_k(z) - tz^{k-1}$ for $t \in [\varepsilon, 1]$. We show that $D_k(z) - tz^{k-1} \neq 0$ for all $z \in \{z \mid |z| = 1\}$. Since $D_k(1) - t < 0$ for $\varepsilon \leq t \leq 1$ it is sufficient to prove that $(z-1)\ \{D_k(z)-tz^{k-1}\} \neq 0$ for all $z \neq 1$ and $|z| = 1$. Suppose $(z-1)\ \{D_k(z)-tz^{k-1}\} = 0$ for some $z \neq 1$ and $|z| = 1$. Then

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

$$\Rightarrow \quad |z^2 - z(s+t) - (1-t)| = s.$$

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

$$2(1-t)y^2 + t(s+t)y - (t^2 + t(s-2) + 2) = 0. \tag{B.4}$$

For $t = 1$, $y = 1$ is the only root and we obtain $\cos\theta = 1$ which contradicts $z \neq 1$. Let $t \in [\varepsilon, 1)$, then the second root of (B.4) is

$$y(t) = \frac{-[t^2+t(s-2)+2]}{2(1-t)} = \frac{-t^2-ts}{2(1-t)} - 1 < -1.$$

Thus we have the contradiction $\cos\theta < -1$ and we get that $D_k(z) - tz^{k-1} \neq 0$ for $z \in \{z \mid |z| = 1\}$.

Observing that for $t = 1$, $D_k(z) - tz^{k-1}$ yields the polynomial $f_k(z)$ with $a = s$, discussed in Theorem 1, we apply Lemma 4 to conclude that for any positive t artibtrarily close to zero the polynomial $D_k(z) - tz^{k-1}$ has k-1 roots <u>inside</u> the disc $\{z \mid |z| \leq 1\}$. Continuity arguments (see for example [6, Appx. A]) lead to the conclusion that $D_k(z)$ has k-1 roots in $\{z \mid |z| \leq 1\}$. By substituting $t = 0$ in (B.4) 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 if and only if k is even. Hence the theorem is proved.

References

[1] D. G. Luenberger, Introduction to Linear and Nonlinear Programming, Addison-Wesley Publishing Co., 1973, Sections 7.2, 7.3.

[2] J. M. Ortega and W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, 1970.

[3] A. Ostrowsky, Solution of Equations and Systems of Equations, Academic Press, New York 1966, 2nd edition.

[4] A. Ralston, "On Differentiating Error Terms," Amer. Math. Monthly, 70, (1963), pp. 187-189.

[5] A. Tamir, "A One-Dimensional Search Based on Interpolating Polynomials Using Function Values Only," to appear in Management Science, 1975.

[6] J. F. Traub, Iterative Methods for Solution of Equations, Prentice-Hall, 1964.