International Series of Numerical Mathematics, Vol. 96, © 1991 Birkhäuser Verlag Basel

AN ENCLOSURE METHOD WITH HIGHER ORDER OF CONVERGENCE-APPLICATIONS TO THE ALGEBRAIC EIGENVALUE PROBLEM

G. Alefeld , University of Karlsruhe

B. Illg , University of Karlsruhe

F. Potra , University of Iowa

Dedicated to L. Collatz on the occassion of his 80th birthday

1. Introduction

In [1] we have considered the nonlinear equation $f(x) = 0$ where f is a continuous differentiable real function of a real variable. We suppose that f is strictly monotone on an interval X^0 . Without loss of generality we may assume that f is strictly increasing on X^0 . We assume that by using interval arithmetic methods it is possible to compute two positive numbers ℓ_1 , ℓ_2 such that $0 < \ell_1 \leq f'(x) \leq \ell_2$ for all $x \in X^0$. Let us denote by L the interval $[\ell_1,\ell_2]$. We suppose that the derivative $f'(x) \in \mathbb{R}$, $x \in X^0$, has an interval extension $f'(X)$, $X \subseteq X^0$, satisfying the following conditions

$$f'(x) \in f'(X), \qquad x \in X \subseteq X^o$$
$$f'(X) \subseteq f'(Y), \qquad X \subseteq Y \subseteq X^o$$
$$d(f'(X)) \leq c\, d(X), \qquad X \subseteq X^o$$

where c is a constant independent of X and where d denotes the diameter of an interval. Furthermore we assume that these three relations also hold for the second derivative of f. Together with f and its derivatives we consider its divided differences

$$f[x,y] = \begin{cases} \dfrac{f(x)-f(y)}{x-y} & \text{if } x \neq y \\ f'(x) & \text{if } x = y \end{cases},$$

$$f[x,y,z] = \begin{cases} \dfrac{f[x,z]-f[y,z]}{x-y} & \text{if } x \neq y \\ \dfrac{f''(x)}{2} & \text{if } x = y \end{cases}.$$

Then for any nonnegative integer p we can define the following iterative procedure.

Algorithm S_p: For $k = 0,1,\ldots$ DO through ES

$x^k = m(X^k)$

if $k = 0$ then $Q^k = L$, $Y^k = X^o$ & GOTO E1

else

$M^k = \{ f[x^k, x^{k-1,p}] + \frac{1}{2} f''(X^{k-1}) (X^k - x^{k-1,p}) \} \cap L$

$Y^k = \{ x^k - f(x^k)/M^k \} \cap Y^k$

$Q^k = \{ f[x^k, x^{k-1,p}] + \frac{1}{2} f''(X^{k-1}) (Y^k - x^{k-1,p}) \} \cap L$

E1 $X^{k,1} = \{ x^k - f(x^k)/Q^k \} \cap X^k$

if $p = 0$ then $x^{k,p} = x^k$ & GOTO ES

else

$x^{k,1} = m(X^{k,1})$

$M^{k,1} = \{ f[x^k, x^{k,1}] + \frac{1}{2} f''(X^k)(X^{k,1} - x^k) \} \cap L$

$Y^{k,1} = \{ x^{k,1} - f(x^{k,1})/M^{k,1} \} \cap X^{k,1}$

$Q^{k,1} = \{ f[x^k, x^{k,1}] + \frac{1}{2} f''(X^k)(Y^{k,1} - x^k) \} \cap L$

$X^{k,2} = \{ x^{k,1} - f(x^{k,1})/Q^{k,1} \} \cap Y^{k,1}$

if $p = 1$ then GOTO ES

else for $i = 2,3,...,p$ DO through E2

$x^{k,i} = m(X^{k,i})$

$M^{k,i} = \{ f[x^{k,i-1}, x^{k,i}] + \frac{1}{2} f''(X^k)(X^{k,i} - x^{k,i-1}) \} \cap L$

E2 $X^{k,i+1} = \{ x^{k,i} - f(x^{k,i})/M^{k,i} \} \cap X^{k,i}$

ES $X^{k+1} = X^{k,p+1}$ □

For S_p we have the following result.

Theorem. Assume that $f(x) = 0$ has a zero x^* in X^o. Moreover assume that the assumptions mentioned before hold. Then the sequence $\{ X^k \}$ generated by S_p is convergent to x^*. Moreover the sequence of diameters $\{ d(X^k) \}$ converges to zero with R-order ω_p defined as

$$\omega_p = (f_p + 2f_{p+2} - 1 + \sqrt{12f_{p+2}^2 + 9f_p^2 - 20f_p f_{p+2} - 4f_{p+2} + 2f_p + 1})/2$$

where f_j denotes the j-th Fibonacci number, i.e.

$$f_0 = 0 \ , \quad f_1 = 1 \ , \quad f_{j+1} = f_j + f_{j-1} \ , \quad j = 1,2,\ldots \ .$$

□

A proof can be found in [1].

Under the assumption that the cost of an interval evaluation of the second derivative is about the same as a function evaluation, the efficiency index of the algorithm in the sense of Ostrowski is given by

$$\mathrm{eff}(S_p) = \sqrt[p+2]{\omega_p} \ .$$

In table I we give the values of ω_p and $(\omega_p)^{1/(p+2)}$ for $p = 0,1,2,\ldots,10$.

Table I. Order and efficiency index of S_p

p	ω_p	$\sqrt[p+2]{\omega_p}$
0	2.00000000000 E+00	1.41421356237 E+00
1	3.73205087570 E+00	1.55113351807 E+00
2	6.46410161514 E+00	1.59450925267 E+00
3	1.10000000000 E+01	1.61539426620 E+00
4	1.82736184955 E+01	1.62294608383 E+00
5	3.00996688705 E+01	1.62638403519 E+00
6	4.92032386541 E+01	1.62741835990 E+00
7	8.01372644808 E+01	1.62756060099 E+00
8	1.30176682947 E+02	1.62724640258 E+00
9	2.11151549918 E+02	1.62677223759 E+00
10	3.42166585524 E+02	1.62624684244 E+00

It can be proved that

$$\omega_p > \left(\frac{1+\sqrt{5}}{2}\right)^{p+2} \quad \text{for } p \geq 4 \ ,$$

$$\lim_{p\to\infty} \sqrt[p+2]{\omega_p} = \frac{1+\sqrt{5}}{2} \ ,$$

$$\max_{p\geq 0} \sqrt[p+2]{\omega_p} = \sqrt[9]{\omega_7} = 1.627... \ .$$

2. The Method for Systems

In the present paper we show how the method repeated in the introduction can be generalized to systems of equations and give some applications to the algebraic eigenvalue problem. For the formulation of the method we need some definitions. Assume that $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is a mapping which has continuous partial derivatives. Then for every pair $x,y \in D$ we define an $n\times n$ matrix $f[x,y]$ by

$$(1) \qquad f[x,y]_{ij} = \begin{cases} \dfrac{1}{x_j - y_j} \{ f_i(x_1,\ldots,x_j,y_{j+1},\ldots,y_n) - \\ \qquad - f_i(x_1,\ldots,x_{j-1},y_j,\ldots,y_n) \} & \text{for } x_j \neq y_j \\ \dfrac{\partial f_i}{\partial x_j}(x_1,\ldots,x_j,y_{j+1},\ldots,y_n) & \text{for } x_j = y_j \ . \end{cases}$$

The matrix $f[x,y]$ is called a "Steigung" or a divided difference operator. It was used by J.W. Schmidt in [3] where the generalization of the Regula falsi to systems of equations was investigated.

For a given mapping $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ and a fixed $z \in D$ we define $\varphi_z : D \subseteq \mathbb{R}^n \to L(\mathbb{R}^n)$ by

$$\varphi_z(x) = f[x,z]$$

and similarly $\psi_x : D \subseteq \mathbb{R}^n \to L(\mathbb{R}^n)$ by

$$\psi_x(z) = f[x,z] .$$

We define the divided difference operator of the second order $f[x,y,z]^1$ by applying (1) to the columns of ψ_x. $f[x,y,z]^3$ is defined similarly. The divided difference operators of the second order are bilinear operators.

For a mapping $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ which has second order partial derivatives with an interval arithmetic evaluation for an interval vector $[u] \subseteq D$, we define three dimensional arrays of intervals $\Delta_s([u]) = (\Delta_s([u])_{ijk})$, $s = 1,2,3$, as follows:

$$\Delta_1([u])_{ijk} = \begin{cases} \frac{1}{2} \frac{\partial^2 f_i([u])}{\partial x_j^2} & \text{for} \quad j = k \\ \frac{\partial^2 f_i([u])}{\partial x_j \partial x_k} & \text{for} \quad k > j \\ 0 & \text{otherwise} \end{cases} ,$$

$$\Delta_2([u])_{ijk} = \begin{cases} \frac{1}{2} \frac{\partial^2 f_i([u])}{\partial x_j^2} & \text{for} \quad j = k \\ 0 & \text{for} \quad k > j \\ \frac{\partial^2 f_i([u])}{\partial x_j \partial x_k} & \text{for} \quad k < j \end{cases} ,$$

$$\Delta_3([u])_{ijk} = \frac{1}{2}\frac{\partial^2 f_i([u])}{\partial x_j \partial x_k} \qquad \text{for} \quad i,j,k = 1(1)n \quad .$$

Using these arrays it can be shown that

$$(f[x,y,z]^1(y-z))_{ij} \in (\Delta_1(x \cup y \cup z)(y-z))_{ij} \ ,$$

$$(f[y,z,x]^3(y-z))_{ij} \in (\Delta_2(x \cup y \cup z)(y-z))_{ij} \ ,$$

$$((f[x,y,z]^1 + f[y,z,x]^3)(y-z))_{ij} \in (\Delta_3(x \cup y \cup z)(y-z))_{ij}$$

where $x \cup y \cup z$ denotes the smallest interval vector containing x, y and z .

Now we set

$$\delta_s(x,u) = \begin{cases} f[x,u] & \text{for} \quad s = 1 \\ f[u,x] & \text{for} \quad s = 2 \\ \frac{1}{2}(f[x,u] + f[u,x]) & \text{for} \quad s = 3 \quad . \end{cases}$$

Assume that the mapping $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ has partial derivatives of second order which can be evaluated in the interval arithmetic sense. Assume that $[L]$ is an interval matrix with $f'(x) \in [L]$ for all $x \in [x]^0$ and that Gaussian elimination can be performed with $[L]$ and an arbitrary interval vector $[u]$. (The result is denoted by IGA $([L],[u])$.) For a given interval vector $[x]$ we denote by $m[x]$ the center of $[x]$. Let $p \geq 0$ be some fixed integer. Then we consider the following iteration methods for $s = 1,2,3$.

Algorithm S_p^s : For $k = 0,1,...$ DO through ES

$x^k = m([x]^k)$

if $k = 0$ then $[Q]^k = [L]$, $[y]^k = [x]^o$ & GOTO E1

else

$[M]^k = \{\delta_s(x^k, x^{k-1,p}) + \Delta_s([x]^{k-1})([x]^k - x^{k-1,p})\} \cap [L]$

$[y]^k = \{x^k - IGA([M]^k, f(x^k))\} \cap [x]^k$

$[Q]^k = \{\delta_s(x^k, x^{k-1,p}) + \Delta_s([x]^{k-1})([y]^k - x^{k-1,p})\} \cap [L]$

E1 $[x]^{k,1} = \{x^k - IGA([Q]^k, f(x^k))\} \cap [y]^k$

if $p = 0$ then $x^{k,p} = x^k$ & GOTO ES

else

$x^{k,1} = m[x]^{k,1}$

$[M]^{k,1} = \{\delta_s(x^{k,1}, x^k) + \Delta_s([x]^k)([x]^{k,1} - x^k)\} \cap [L]$

$[y]^{k,1} = \{x^{k,1} - IGA([M]^{k,1}, f(x^{k,1}))\} \cap [x]^{k,1}$

$[Q]^{k,1} = \{\delta_s(x^{k,1}, x^k) + \Delta_s([x]^k)([y]^{k,1} - x^k)\} \cap [L]$

$[x]^{k,2} = \{x^{k,1} - IGA([Q]^{k,1}, f(x^{k,1}))\} \cap [y]^{k,1}$

if $p = 1$ then GOTO ES else

for $i = 2,3,...,p$ DO through E2

$x^{k,i} = m[x]^{k,i}$

$[M]^{k,i} = \{\delta_s(x^{k,i}, x^{k,i-1}) + \Delta_s([x]^k)([x]^{k,i} - x^{k,i-1})\} \cap [L]$

E2 $[x]^{k,i+1} = \{x^{k,i} - IGA([M]^{k,i}, f(x^{k,i}))\} \cap [x]^{k,i}$

ES $[x]^{k+1} = [x]^{k,p+1}$ □

If $f(x^*) = 0$ for some $x^* \in [x]^0$ (under our assumptions x^* is unique) then $x^* \in [x]^k$, $k \geq 0$.

Furthermore it has been proved under appropriate assumptions that the R-order of convergence of S_p^s is the same as in the one dimensional case.

4. Applications

In general the method S_p^s seems not to be very attractive for systems in n unknowns since one needs approximately n^3 interval arithmetic evaluations for the second order partial derivatives. However, there are some important cases in which one needs less work.

a) Consider the nonlinear integral equation

$$\int_0^1 K(t,s,x(t))dt = x(s) , \quad s \in [0,1]$$

for the unknown function $x(s)$. For the numerical solution of this equation we choose equidistant points $s_i = \frac{i}{n}$, $i = 0(1)n$, and use one of the well known numerical integration formulas. Omitting the discretization error, we get a nonlinear system $f(x) = 0$ with

$$f_i(x_0,\ldots,x_n) = x_i - \sum_{j=0}^{n} w_j K(\tfrac{j}{n},\tfrac{i}{n},x_j) , \quad i = 0(1)n ,$$

for the unknowns x_i , $i = 0(1)n$, which are considered as approximations to $x(\frac{i}{n})$, $i = 0(1)n$. It follows that

$$\frac{\partial f_i}{\partial x_j} = \delta_{ij} - w_j K_u(\frac{j}{n},\frac{i}{n},x_j) \quad , \quad i = 0(1)n \quad , \quad j = 0(1)n$$

and

$$\frac{\partial^2 f_i}{\partial x_j \partial x_k} = \begin{cases} - w_j K_{uu}(\frac{j}{n},\frac{i}{n},x_j) & \text{for} \quad j = k \\ 0 & \text{otherwise} \end{cases} \quad , \; i = 0(1)n \, , \; j = 0(1)n \, , \; k = 0(1)n \, .$$

(δ_{ij} denotes the Landau-symbol, K_u and K_{uu} denote the first and second order partial derivative with respect to the third variable of K). Hence in this case f'' has only $(n+1)^2$ elements different from zero.

b) A similar result as in a) holds if a solution of the boundary value problem

$$y'' = f(t,y)$$
$$y(a) = \alpha \, , \quad y(b) = \beta$$

is approximated by the usual method of differences.

c) Even more spectacular than in the two proceeding examples is the saving of arithmetic operations for the algebraic eigenvalue problem. Consider the eigenvalue problem for the matrix A . If we define the vector $x = (z^T,\lambda)^T$ then

$$A z = \lambda z$$
$$z^T z = 1$$

is equivalent to the nonlinear system

$$f(x) = 0$$

where

$$f_i(x) = (a_{ii} - x_{n+1})\, x_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij}\, x_j \quad , \qquad i = 1\,(1)\,n$$

$$f_{n+1}(x) = \sum_{i=1}^{n} z_i^2 - 1 \quad .$$

In this case we get

$$\frac{\partial f_i(x)}{\partial x_j} = \begin{cases} a_{ij} - \delta_{ij}\, x_{n+1} & , \quad 1 \le i \le n \quad , \; 1 \le j \le n \\ -x_i & , \quad 1 \le i \le n \quad , \; j = n+1 \\ 2\, x_j & , \quad i = n+1 \quad , \; 1 \le j \le n \\ 0 & , \quad i = j = n+1 \end{cases} ,$$

$$\frac{\partial^2 f_i(x)}{\partial x_j\, \partial x_k} = \begin{cases} 0 & , \quad 1 \le i \le n \quad , \; 1 \le j \le n \; , \; 1 \le k \le n \\ -\delta_{ij} & , \quad 1 \le i \le n \quad , \; 1 \le j \le n \; , \; k = n+1 \\ -\delta_{ik} & , \quad 1 \le i \le n \quad , \; j = n+1 \quad , \; 1 \le k \le n+1 \\ 2\,\delta_{jk} & , \quad i = n+1 \quad , \; 1 \le j \le n \; , \; 1 \le k \le n+1 \\ 0 & , \quad i = j = n+1 \quad , \; 1 \le k \le n+1 \end{cases} ,$$

$$f[x,y]_{ij} = \begin{cases} a_{ij} - \delta_{ij}\, y_{n+1} & , \quad 1 \le i \le n \quad , \; 1 \le j \le n \\ -x_i & , \quad 1 \le i \le n \quad , \; j = n+1 \\ x_j + y_j & , \quad i = n+1 \quad , \; 1 \le j \le n \\ 0 & , \quad i = j = n+1 \end{cases} .$$

Hence the second derivative is constant and many of its elements are equal to zero. Similarly f[x,y] can be formed nearly without arithmetical work. Hence S_p^s can be performed with simple available operators.

4. Numerical Examples

a) The matrix

$$A = \begin{bmatrix} 33 & 16 & 72 \\ -24 & -10 & -57 \\ -8 & -4 & -17 \end{bmatrix}$$

has an eigenpair $x = (z^T, \lambda)^T$ which is contained in the intervalvector

$$[x]^0 = \begin{bmatrix} [-0.765, -0.764] \\ [0.611, 0.612] \\ [0.203, 0.204] \\ [0.991, 1.001] \end{bmatrix} .$$

The following table II contains the numerical results obtained by applying S_p^3 for different values of p. For a fixed p the integer k denotes the number of iteration steps until the lower and upper bounds of the iterates $[x]^k$ differ by at most one unit of the last digit in the mantissa. (We are using a computer with 12 decimal digits in the mantissa.) f denotes the number of function evaluations and IGA is the number of applications of Gaussian elimination.

Table II			
p	k	f	IGA
0	2	3	4
1	1	3	4
2	1	3	4
3	0	3	4
4	0	3	4
5	0	3	4
6	0	3	4

b) The matrix

$$A = \begin{bmatrix} -2 & 1 & 0 & 27 & -18 & -6 \\ -8 & 4 & 0 & 54 & -36 & -12 \\ -8 & -5 & 6 & 81 & -54 & -18 \\ -8 & -5 & -6 & 117 & -72 & -24 \\ -8 & -5 & -6 & 129 & -78 & -30 \\ -8 & -5 & -6 & 129 & -60 & -48 \end{bmatrix}$$

has an eigenpair $x = (z^T, \lambda)^T$, which is contained in the intervalvector

$$[x]^0 = \begin{bmatrix} [\ 0.127\ ,\ 0.128\ ] \\ [\ 0.254\ ,\ 0.255\ ] \\ \vdots \\ [\ 0.508\ ,\ 0.509\ ] \\ [11.991\ ,\ 12.01\ ] \end{bmatrix} .$$

The values in table III have the analogous meaning as in table II .

Table III

p	k	f	IGA
0	2	3	4
1	1	3	4
2	1	3	4
3	0	3	4
4	0	3	4
5	0	3	4
6	0	3	4
7	0	3	4

References

[1] Alefeld, G., Potra, F.: A new class of interval methods with higher order of convergence. Computing 42, 69-80 (1989).

[2] Illg, B.: Über einige Verfahren höherer Ordnung zur iterativen Einschließung bei nichtlinearen Gleichungssystemen. Diplomarbeit. Universität Karlsruhe 1989. (Not available).

[3] Schmidt, J. W.: Eine Übertragung der Regula falsi auf Gleichungen in Banachräumen II. ZAMM 43, 97-110 (1963).

G. Alefeld , B. Illg
Institut für Angewandte Mathematik
Universität Karlsruhe
Kaiserstraße 12
D-7500 Karlsruhe
Deutschland

F. Potra
Department of Mathematics
University of Iowa
Iowa City , 117 522 42
U. S. A.