THE OPTIMAL EFFICIENCY INDEX OF A CLASS OF HERMITE ITERATIVE METHODS WITH TWO STEPS

ION PĂVĂLOIU

Abstract. The inverse interpolatory polynomials of Hermite type with 2 nodes, all having the same order of multiplicity $q \in \mathbb{N}$, provide a class of iterative methods for solving scalar equations. In this note we determine the iterative method with the highest efficiency index: the optimal method is obtained for $q = 2$.

1. INTRODUCTION

As we have shown in [7], the problem of determining the Hermite interpolatory interative methods having the optimal efficiency index cannot be solved in the general case. We have determined however in [7] some upper and lower bounds for the efficiency indexes; these bounds depend on the coefficients of the characteristic equation whose positive root provide the convergence order of the considered equation.

In this note we shall consider a subclass of Hermite interpolatory methods, based on two interpolatory nodes which have the same multiplicity order $q \in \mathbb{N}$. We shall use the efficiency index defined in [4] for determining the method with the optimal efficiency index.

Let $I = [a, b]$, $a, b \in R$, $a < b$, $f : I \to R$ and consider the equation

$$(1) \qquad f(x) = 0.$$

We assume that this equation has a unique solution $\overline{x} \in]a, b[$. For solving the above equation we consider a function $g : I \to I$ and we also assume that $\overline{x}$ is the unique fixed point of g in I.

For approximating $\overline{x}$ there is usually taken an element from the sequence $(x_s)_{s \geq 0}$ generated by the iterative method:

$$(2) \qquad x_{s+1} = g(x_s), \quad s = 0, 1, ..., \quad x_0 \in I.$$

More generally, if $h : I^k \to I$ is a function of k variables whose restriction to the diagonal of the set I^k coincides with g, then the following sequence may

---

be considered for approximating $\overline{x}$:

(3) $\qquad x_{s+k} = h\left(x_s, x_{s+1}, ..., x_{s+k-1}\right), \quad s = 0, 1, ..., \; x_0, x_1, ..., x_{k-1} \in I.$

The convergence of the sequence $(x_s)_{s \geq 0}$ generated by (2) or (3) depends on certain properties of the functions f, g and resp. h. The time needed by a computer to obtain a convenient approximation of $\overline{x}$ depends on the convergence order of the sequence $(x_s)_{s \geq 0}$ and also on the number of elementary operations performed at each iteration step.

While the convergence order may be determined exactly in most of the situations, the number of elementary operations may be hard to evaluate.

For this reason Ostrowski has proposed in [4] a simplification of this problem, by considering the number of function evaluations needed at each iteration step. This leads to the definition of the efficiency index, which may be naturally applied to the comparison of different methods.

Consider a sequence $(x_s)_{s \geq 0}$ which together with f and g satisfies:

    a) $x_s \in I$; for all $s = 0, 1, ...$

    b) $(x_s)_{s \geq 0}$ converges to $\overline{x}$;

    c) the sequence $g\left((x_s)\right)_{s \geq 0}$ converges also to $\overline{x}$;

    d) $f(\overline{x}) = 0$;

    e) f is derivable at $\overline{x}$;

    f) for all $x, y \in I$ there exists $m \in R_+$, such that $0 < [x, y; f] < m$, where $[x, y; f]$ denotes the divided differences of f on the nodes x and y.

Concerning the convergence order of $(x_s)_{s \geq 0}$, we shall consider the following definition:

DEFINITION 1. [4], *The sequence $(x_s)_{s \geq 0}$ has the convergence order $\omega \in R$, $\omega \geq 1$, with respect to the function g if there exists the limit*

$$\alpha = \lim_{s \to \infty} \frac{\ln |g(x_s) - \overline{x}|}{\ln |x_s - \overline{x}|}$$

and $\alpha = \omega$.

LEMMA 2. [7], *If the sequence $(x_s)_{s \geq 0}$ and the functions f and g satisfy conditions a) – f), then the necessary and sufficient condition for the sequence $(x_s)_{s \geq 0}$ to have the convergence order $\omega \in R$, $\omega \geq 1$, with respect to g is that there exists the limit*

$$\beta = \lim_{s \to \infty} \frac{\ln |f(g(x_s))|}{\ln |f(x_s)|}$$

and $\beta = \omega$.

LEMMA 3. [7], *If $(u_s)_{s \geq 0}$ is a sequence of real positive numbers satisfying*

    i) *the sequence $(u_s)_{s \geq 0}$ converges to 0;*

    ii) *there exist the nonnegative real numbers $\alpha_1, \alpha_2, ..., \alpha_{n+1}$ and a bounded sequence of real positive numbers $(c_s)_{s \geq 0}$, $0 < \inf_s \{c_s\}$ and the following equalities hold:*

$$u_{s+n+1} = c_s u_s^{\alpha_1} u_s^{\alpha_1} \cdots u_s^{\alpha_1}, \quad s = 0, 1, ...;$$

    iii) *the sequence $\left(\dfrac{\ln u_{s+1}}{\ln u_s}\right)_{s \geq 0}$ is convergent and $\lim \dfrac{\ln u_{s+1}}{\ln u_s} = \omega > 0$, then ω is the positive root of the equation*

$$t^{n+1} - \alpha_{n+1} t^n - \alpha_n t^{n-1} - \cdots - \alpha_2 t - \alpha_1 = 0.$$

We shall denote in the following by m_s the number of function evaluations that must be performed at the step s for the iteration (2) or (3).

Taking into account Lemmas 1 and 2, the following definition becomes natural:

DEFINITION 4. [4], *The real number E is called the efficiency index of the method (2) (resp. (3)) if there exists*

$$l = \lim \left(\frac{\ln |f(x_{s+1})|}{\ln |f(x_s)|} \right)^{\frac{1}{m_s}}$$

and $l = E$.

Remark. If there exists $s_0 \in \mathbb{N}$ such that for $s > s_0$ the values m_s are constant, $m_s = r$, then the efficiency index is given by relation

(4) $\qquad\qquad\qquad E = \omega^{\frac{1}{r}}.$

2. TWO STEP HERMITE INTERPOLATORY ITERATIVE METHODS

Let $q \in \mathbb{N}$, $q \geq 1$ be a natural number and consider the Hermite inverse interpolatory polynomial on two nodes with the same multiplicity order q.

We shall make the following assumptions concerning the function f:

    α) f is derivable on $]a, b[$ up to and including the $2q$-th order;

    β) $f'(x) \neq 0$ for all $x \in]a, b[$;

    γ) equation (1) has a solution $\overline{x} \in]a, b[$.

Under these assumptions, it is clear that f admits an inverse $f^{-1} : D \to I$, where $D = f(I)$, and also that $\overline{x}$ is given by the relation

$$\overline{x} = f^{-1}(0).$$

Moreover, f^{-1} is derivable up to the order $2q$ at all points from D and the k-th derivative, for $k \in \mathbb{N}$, $1 \leq k \leq 2q$ is given by [9].

(5)

$$\left[f^{-1}(y)\right]^{(k)} = \sum \frac{(2k - i_1 - 2)! \, (-1)^{k+i_1-1}}{i_2! i_3! ... i_k! \, [f'(x)]^{2k-1}} \left(\frac{f'(x)}{1!}\right)^{i_1} \cdots \left(\frac{f^{(k)}(x)}{k!}\right)^{i_k},$$

where we have denoted $y = f(x)$, and the above sum extends on the nonnegative whole solutions of the system

$$i_2 + 2i_3 + ... + (k-1)i_k = k - 1$$
$$i_1 + i_2 + ... + i_k = k - 1$$

Denote by $H(y_1, q; y_2, q; f^{-1}|y)$ the Hermite inverse interpolatory polynomial satisfying

$$(6) \quad H^{(k)}(y_1, q; y_2, q; f^{-1}|y_i) = [f^{-1}(y_i)]^{(k)}, \; i = 1, 2; \; k = 0, 1, ..., q-1,$$

where $[f^{-1}(y_i)]^{(0)} = f^{-1}(y_i)$, $i = 1, 2$ and $y_1, y_2 \in D$.

Consider also the polynomial

$$\omega_1(y) = (y - y_1)^q (y - y_2)^q.$$

The residual in the interpolation formula becomes then

$$R(f^{-1}; y) = f^{-1}(y) - H(y_1, q; y_2, q; f^{-1}|y) =$$
$$= \frac{1}{(2q)!}[f^{-1}(\theta_1)]^{(2q)} \omega_1(y),$$

where θ_1 belongs to the smallest open interval determined by the points y, y_1 and y_2.

Let $x_s, x_{s+1} \in I$ be two approximations of the solution $\bar{x}$. Then the next approximation may be determined by

$$(7) \quad x_{s+2} = H(y_s, q; y_{s+1}, q; f^{-1}|0), \; s = 0, 1..., .$$

We shall assume that all the elements of the sequence $(x_s)_{s \geq 0}$ generated by (7) belong to the interval I.

Taking into account the above assumptions we easily get for $s = 0, 1, ...$ that

$$|f(x_{s+2})| = |f'(\alpha_s)| \frac{\left|[f^{-1}(\theta_s)]^{(2q)}\right|}{(2q)!}|f(x_{s+1})|^q |f(x_s)|^q, \; s = 0, 1, ...,$$

where α_s belongs to the open interval determined by the points $\bar{x}$ and x_{s+2}, and θ_s is contained in the open interval determined by $0, y_s, y_{s+1}$.

Assuming that $[f^{-1}(y)]^{(2q)} \neq 0$ for all $y \in intD$, denoting

$$c_s = |f'(\alpha_s)| \frac{\left|[f^{-1}(\theta_s)]^{(2q)}\right|}{(2q)!}, \; s = 0, 1, ...$$

and applying Lemma 2, we obtain the following equation for determining the convergence order of method (7):

$$t^2 - qt - q = 0.$$

It follows that the convergence order is

$$\omega = \frac{q + \sqrt{q^2 + 4q}}{2}.$$

3. THE OPTIMAL EFFICIENCY INDEX

We remark that in order to generate the elements of the sequence $(x_s)_{s \geq 0}$ given by (7), at each iteration step s, $s \geq 2$, the following function evaluations are needed:

$$f, f', ..., f^{(q-1)},$$

all at x_s, since their values at x_{s-1} are known from the previous step. Hence there are needed q function evaluations.

Remark. Taking into account that the Hermite inverse interpolatory polynomial is computed with the aid of the succession derivatives of f^{-1}, which, by (5) have a rather complicated form, then it becomes necessary to take into account $q - 1$ more function evaluations. On the other hand, the evaluation of the Hermite polynomial determined by (6) may lead us to the consideration of a one more function evaluation. We can therefore conclude that at each iteration step there is necessary an amount of $2q$ function evaluations. These considerations may affect the value of the efficiency index, but we shall see in the following that the optimal efficiency index is not affected by considering a number of function evaluations proportional with q.

Indeed, considering the number of function evaluations as being equal to δq, δ being a positive constant, then by (4), the efficiency index of method (7) is

$$(8) \quad E = \varphi(q) = \left[\frac{q + \sqrt{q^2 + 4q}}{2}\right]^{\frac{1}{\delta q}}.$$

The value of q for which E attains the upper bound is given by the solution of $\varphi'(q) = 0$, and we see below that this solution does not depend on δ.

By (8) we get

$$\varphi'(q) = \frac{1}{\delta}\varphi(q)\left[\frac{1}{q}\ln\frac{q + \sqrt{q^2 + 4q}}{2}\right]'.$$

Since $\varphi(q) > 0$, it follows that equation $\varphi'(q) = 0$ is equivalent with the following one

$$\left(\frac{1}{q}\ln\frac{q + \sqrt{q^2 + 4q}}{2}\right)' = 0,$$

whence we get

$$(9) \qquad \Psi(q) = \frac{\sqrt{q^2 + 4q} + q + 2}{\sqrt{q^2 + 4q} + q + 4} - \ln \frac{q + \sqrt{q^2 + 4q}}{2} = 0.$$

For solving this equation denote $t = q + \sqrt{q^2 + 4q}$ and we notice that $\frac{dt}{dq} > 0$ for $q > 0$. By this substitution, equation (9) becomes

$$\eta(t) = \frac{t + 2}{t + 4} - \ln \frac{t}{2} = 0,$$

and since $\eta'(t) < 0$ for $t > 0$, it follows that equation $\eta(u) = 0$ has a unique positive solution $\bar{t}$. We also notice that $\eta(2) = \frac{2}{3} > 0$ and $\eta(2e) = \frac{e+1}{e+2} - 1 < 0$, i.e. $2 < E < 2e$, which lead us to the conclusion that the positive root $\bar{q}$ of equation $\varphi'(q) = 0$ satisfies

$$2 < \bar{q} + \sqrt{\bar{q}^2 + 4\bar{q}} < 2e,$$

and whence

$$(10) \qquad \frac{1}{2} < \bar{q} < \frac{e^2}{e + 1}.$$

We also remark that $\eta(t) > 0$ for $1 \le t \le \bar{t}$ and $\eta(t) < 0$ for $\bar{t} < t$, so it follows that $\varphi'(q) > 0$ for $\frac{1}{2} < q < \bar{q}$ and $\varphi'(q) < 0$ for $q > \bar{q}$. The function $E = \varphi(q)$ has at $q = \bar{q}$ a maximum value. It remains to compute the maximum value of E in the set of natural numbers from the neighborhood of the real number $\bar{q}$. By (10), the value of q for which E attains the maximum belongs to the set $\{1, 2, 3\}$. One can easily check that $\varphi(1) < \varphi(2)$ and $\varphi(2) > \varphi(3)$, which implies that E attains the maximum value for $q = 2$.

We have proved the following theorem:

THEOREME 5. *Among all the iterative methods (7), the method with the highest efficiency index is the one corresponding to $q = 2$, being given by*

$$(11) \qquad x_{s+2} = H\left(y_s, 2; y_{s+1}, 2; f^{-1} | 0\right), x_0, x_1 \in I, \ s = 0, 1, ..., .$$

Finally we give for H an expression based on the divided differences on double nodes (see [6]).

$$(12) \qquad \begin{aligned} H\left(y_s, 2; y_{s+1}, 2; f^{-1} | 0\right) &= x_s - \left[y_s, y_s; f^{-1}\right] y_s + \\ &+ \left[y_s, y_s, y_{s+1}; f^{-1}\right] y_s^2 - \\ &- \left[y_s, y_s, y_{s+1}, y_{s+1}; f^{-1}\right] y_s^2 y_{s+1}, \end{aligned}$$

where $y_s = f(x_s)$, $y_{s+1} = f(x_{s+1})$.

For computing the divided differences from (12) one can use the recurrence formula for divided differences with the aid of the following table:

y_s	$f^{-1}(y_s)$			
y_s	$f^{-1}(y_s)$	$[y_s, y_s; f^{-1}]$		
y_{s+1}	$f^{-1}(y_{s+1})$	$[y_s, y_{s+1}; f^{-1}]$	$[y_s, y_s, y_{s+1}; f^{-1}]$	
y_{s+1}	$f^{-1}(y_{s+1})$	$[y_{s+1}, y_{s+1}; f^{-1}]$	$[y_s, y_{s+1}, y_{s+1}; f^{-1}]$	$[y_s, y_s, y_{s+1}, y_{s+1}; f^{-1}]$

where

$$y_s = f(x_s), \ y_{s+1} = f(x_{s+1}), \ \left[y_s, y_s; f^{-1}\right] = \frac{1}{f'(x_s)},$$

$$\left[y_s, y_{s+1}; f^{-1}\right] = \frac{1}{[x_s, x_{s+1}; f]}$$

and

$$\left[y_{s+1}, y_{s+1}; f^{-1}\right] = \frac{1}{f'(x_{s+1})}.$$

REFERENCES

[1] BREENT, R., WINOGRAD, S., WOLFE, F., *Optimal Iterative Processes for Root-Finding*, Numer. Math., **20**, pp. 327–341, 1973.

[2] COMAN, GH., *Some Practical Approximation Methods for nonlinear Equations*, Matematica-Revue d'Analyse Numérique et de Théorie de l'Approximation, **11**, no. 1–2, pp. 11–48, 1982.

[3] KUNG, H. T., TRAUB., J. F., *Optimal Order and Efficiency for Iterations with Two Evaluations*, SIAM J. Numer. Anal., **13**, no. 1, pp. 84–99, 1976.

[4] OSTROWSKI, A. M., *Solution of Equation and Systems of Equations*, Academic Press, New York and London, 1966.

[5] PĂVĂLOIU, I., *Bilateral Approximations for the Solutions of Scalar Equations*, Revue D'Analyse Numérique et de Théorie de l'Approximation, 23, 1, pp. 95–100, 1994.

[6] PĂVĂLOIU, I., *Optimal Problems Concerning Interpolation Methods of Solutions of Equations*, Publ. Inst. Math, **52** (**66**), pp. 113–126, 1992.

[7] PĂVĂLOIU, I., *Optimal Efficiency Index for Iterative Methods of Interpolatory Type*, Computer Science Journal of Moldova, **5**, nr. 1 (13), pp. 20–43, 1997.

[8] TRAUB, J. F., *Iterative methods for solution of equations*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1964.

[9] TUROWICZ, B. A., *Sur les dérivées d'ordre superieur d'une fonction inverse*, Ann. Polon. Math., **8**, pp. 265–269, 1960.