CAN ANY STATIONARY ITERATION USING LINEAR INFORMATION BE GLOBALLY CONVERGENT?

G.W. WASILKOWSKI

CUCS-143-80
Can Any Stationary Iteration Using Linear Information Be Globally Convergent?

G. W. WASKILOWSKI

Carnegie Mellon University, Pittsburgh, Pennsylvania

ABSTRACT All known globally convergent iterations for the solution of a nonlinear operator equation \( f(x) = 0 \) are either nonstationary or use nonlinear information. It is asked whether there exists a globally convergent stationary iteration which uses linear information. It is proved that even if global convergence is defined in a weak sense, there exists no such iteration for as simple a class of problems as the set of all analytic complex functions having only simple zeros. It is conjectured that even for the class of all real polynomials which have real simple zeros there does not exist a globally convergent stationary iteration using linear information.

KEY WORDS AND PHRASES nonlinear equations, linear information, stationary iterations, global convergence

CR CATEGORIES 5.15

1. Introduction

Suppose we solve a nonlinear operator equation \( f(x) = 0 \) by an iteration which constructs a sequence of approximations \( (x_i) \). For most convergence theorems it is assumed that \( f \) is sufficiently smooth and starting points are "sufficiently close" to a solution \( x \). In practice it is very hard to verify the second assumption. One therefore wants to use iterations which are globally convergent.

All known globally convergent iterations are either nonstationary or use nonlinear information. For instance, Laguerre's iteration (see, e.g., Ralston and Rabinowitz [1]) is globally convergent for the class of all real polynomials having only real zeros. However this iteration uses the degree of the polynomial whose zero is approximated. This means that Laguerre's iteration uses nonlinear information (see Section 6). An example of a nonstationary iteration using linear information which is globally convergent for analytic operator equations may be adopted from Traub and Wozniakowski [4]. This will be reported in Waskiowski [6].

However, most commonly used iterations are stationary and use linear information. Therefore it is important to know whether there exist globally convergent stationary iterations which use linear information. In this paper we prove that for as simple a class of problems as the set of all analytic complex functions with simple zeros, there exists no such iteration. We conjecture that the same negative result holds even for the class of all real polynomials having real simple zeros.

We summarize the contents of the paper. In Section 2 we remind the reader of the definitions of information and stationary iteration without memory. In Section 3 we discuss...
the concept of global convergence. In Section 4 we show that no stationary iteration without memory which uses linear information can be globally convergent for the class of all analytic complex functions. In Section 5 we extend this result to all stationary iterations with or without memory using linear information. In Section 6 we pose a conjecture that for the class of all real polynomials with real simple zeros, there does not exist a globally convergent stationary iteration using linear information.

2. Stationary Iterations Without Memory

We recall the definition of information and stationary iteration. (See Traub and Woźniakowski [5].) For the reader's convenience, in Sections 2 to 4 we deal only with one-point iterations without memory. The extension to the general case is given in Section 5.

Consider the solution of a nonlinear scalar equation

\[ f(x) = 0 \]  \hspace{1cm} (2.1)

for \( f \in \mathcal{F} \), where \( \mathcal{F} \) is a subset of a space \( \mathcal{H} \) of functions \( f : D_f \subset \mathcal{C} \to \mathcal{C} \). To solve (2.1) iteratively, we need to know something about \( f \). Let \( L_i : D_{L_i} \subset \mathcal{H} \times \mathcal{C} \to \mathcal{C} \) be a functional which is linear with respect to the first argument, i.e.,

\[ L_i(c_1 f_1 + c_2 f_2, x) = c_1 L_i(f_1, x) + c_2 L_i(f_2, x) \quad \text{whenever} \quad x \in D_{L_i} \cap D_{L_j}, \quad i, j = 1, 2, \ldots, n. \]

Consider the linear information operator \( \mathcal{L} : D_f \subset \mathcal{H} \times \mathcal{C} \to \mathcal{C}^n \), defined as

\[ \mathcal{L}(f, x) = \begin{cases} [L_1(f(x)), L_2(f(x)), \ldots, L_n(f(x))] & \text{for} \quad x \in D_{L_i}, \\ \text{undefined} & \text{otherwise}. \end{cases} \]  \hspace{1cm} (2.2)

for every \( f \in \mathcal{H} \).

Let \( x_0 \) be an approximation of a solution of (2.1). Let \( \varphi : D_\varphi \subset \mathcal{C}^{n+1} \to \mathcal{C} \) be a functional. We construct the sequence of approximations \( x \), by the formula

\[ x_{i+1} = \varphi(x_i, \mathcal{L}(f, x_i)). \]  \hspace{1cm} (2.3)

The functional \( \varphi \) is called a one-point stationary iterative operator without memory using a linear information operator \( \mathcal{L} \). For brevity \( \varphi \) is called an iteration. Let \( \Phi(\mathcal{L}) \) be the class of all such iterations.

Note that most iterations use values of \( f \) and its derivatives. A linear information operator is a generalization of this. For example, the information operator \( g \) used by Newton iteration,

\[ x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}, \]

is \( \mathcal{L}(f, x) = [f(x), f'(x)] \). This operator is linear and \( \varphi \in \Phi(\mathcal{L}) \).

3. What Do We Mean by Global Convergence?

Let \( \mathcal{F} \) be the class of all functions \( f, f \in D_f \subset \mathcal{C} \). For \( f \) analytic in \( D_f \) and having only simple zeros. Let \( S(f) \) be the set of all zeros of \( f \).

Consider any iteration \( \varphi \in \Phi \), where \( \Phi \) is the class of all stationary iterations without memory which use a linear information operator \( g \) per iterative step, i.e., \( \Phi = \{ \Phi(\mathcal{L}) \} \).

Which properties should \( \varphi \) have to be called a globally convergent iteration? To motivate our definition, consider first the problem \( f(x) = 0 \) where \( f \) is defined in \( D_f \subset \mathcal{C} \) and \( a \) is its unique simple zero. Suppose we apply \( \varphi \) to this problem. Let \( r(f, \varphi) \) be the maximal number such that for any starting point \( x_0 \) satisfying \( |x_0 - a| < r(f, \varphi) \), the sequence \( x_{i+1} = \varphi(x_i, \mathcal{L}(f, x_i)) \) is convergent to \( a \). Then the ball \( \{ x : |x - a| < r(f, \varphi) \} \) is called the ball of convergence of \( \varphi \) for \( f \). Of course, \( r(f, \varphi) \) depends on \( R(f) \) and, in particular, \( r(f, \varphi) \leq R(f) \). Suppose there exists a positive constant \( c = c(\varphi) \) such that \( r(f, \varphi) \geq cR(f) \) for any \( f \). Then \( \varphi \) enjoys a type of global convergence, since
the ball of convergence has radius proportional to the radius $R(f)$ of the domain $D$. However, for problems with $R(f) = +\infty$, we get $r(f, q) = +\infty$, which means that any choice of a starting point $x_0$ yields convergence. This seems to be too strong. For $R(f) = +\infty$ we would like to have $r(f, q)$ large but not necessarily equal to infinity. This is the motivation of the following: Let $K, L$ be two given constants such that $K \geq 0$ and $0 < L \leq +\infty$. Define $R_f = \min \{ R(f), K|a| + L \}$. Now the existence of a positive constant $c = c(f)$ such that $r(f, q) \geq cR_f$ implies a type of global convergence of $q$. This discussion shows that we should compare $r(f, q)$ with $R_f$.

If $f$ has more than one simple zero, we proceed as follows. Let $\text{dist}(a, \partial D_f)$ be the distance of $a, a \in S(f)$, to the boundary $\partial D_f$ of the domain $D_f$. Define

$$R_f(a) = \min \{ \text{dist}(a, \partial D_f), K|a| + L \} \quad (K \geq 0, 0 < L \leq +\infty).$$

Note that if $S(f) = \{ a \}$, then $R_f(a) = R_f$. Let

$$B(b, f) = \bigcup_{a \in S(f)} \{ x : |x - a| < b \cdot R_f(a) \}.$$

where $b \geq 0$. For any iteration $q, \varphi \in \Phi(\varphi)$, we define a number $c(\varphi)$ such that

(i) for any $f \in F$ and any starting point $x_0 \in B(c(\varphi), f)$, the sequence $(x_i, x_{i+1}) = \varphi(x_i, I(f, x_i))$, is well defined and $\lim_{i \to \infty} x_i \in S(f)$; and

(ii) for any $\epsilon > 0$, there exist a problem $f \in F$ and a starting point $x_0 \in B(c(\varphi) + \epsilon, f)$ such that either the sequence $(x_i)$ is not well defined or $\lim_{i \to \infty} x_i \notin S(f)$.

Note that for any iteration, $c(\varphi) \in [0, 1]$. The set $B(c(\varphi), f)$ is a convergence domain of $\varphi$ for the function $f$, since taking any starting point $x_0 \in B(c(\varphi), f)$ we get convergence of $(x_i)$. Note, however, that we do not specify which element from $S(f)$ is the limit of $(x_i)$.

**Definition 3.1.** We shall say that an iteration $q, \varphi \in \Phi$, is globally convergent for the class $F$ iff $c(\varphi) > 0$. □

Definition 3.1 imposes only a weak condition on $q$. However, we shall show that for any iteration from $\Phi$, $c(\varphi) = 0$. This means that even in the sense of Definition 3.1 there exist no globally convergent stationary iteration using linear information for the class $F$.

4. Main Result

**Theorem 4.1.** No iteration $q$ from $\Phi$ is globally convergent for the class $F$.

**Proof.** Suppose there exist $q : D \to C^n$ and $q \in \Phi(\varphi)$ with $c = c(q) > 0$. Let

$$a = \begin{cases} 1 & \text{if } cK \geq 1, \\ \frac{cL}{2(1 - cK)} & \text{if } cK < 1. \end{cases}$$

Define $f(x) = x - a$. Then $f \in F$, $S(f) = \{ a \}$, and $R_f(a) = K|a| + L$. Let $x_0 = 0$. Since $x_0 \in B(c,f) = \{ x : |x - a| \leq cR_f(a) \}$, the sequence $(x_i, x_{i+1}) = \varphi(x_i, I(f, x_i))$, tends to $a$. Thus there exists a unique integer $k, k \geq 1$, such that

$$|x_0|, |x_1|, \ldots, |x_k| < \frac{a}{2} \quad \text{and} \quad |x_k| \geq \frac{a}{2} \quad (4.1)$$

Consider a polynomial $w$ of the form

$$w(x) = \sum_{i=0}^{a} a_i x^{i+1}$$

which satisfies

$$s(f - w, x_j) = s(f, x_j) \quad \text{for } j = 0, 1, \ldots, k - 1 \quad (4.3)$$
This is equivalent to the following homogeneous system of \( n \cdot k \) linear equations:

\[
\sum_{j=0}^{n} a_j L_j(x^{s-1}, x) = 0 \quad \text{for} \quad s = 1, 2, \ldots, n \quad \text{and} \quad j = 0, 1, \ldots, k - 1
\]  

(4.4)

Since we have more unknowns than equations, it is obvious that there exists a nonzero polynomial \( w \) of the form (4.2). Then there exist positive integers \( r, 1 \leq r \leq nk + 1 \), and \( p, 1 \leq p \leq r \), and nonzero \( z_{p+1}, \ldots, z \), such that

\[ w(x) = x^p(x - z_{p+1}) \cdots (x - z) \]

For \( \varepsilon > 0 \), define

\[
f_\varepsilon(x) = \begin{cases} 
  f(x) + \frac{1}{\varepsilon} w(x) & \text{if } |x| < \frac{a}{2} \\
  \text{undefined} & \text{if } |x| \geq \frac{a}{2}
\end{cases}
\]

From the general theory of algebraic functions (see, e.g., Wilkinson [7]), it is known that for sufficiently small \( \varepsilon \), there exist \( p \) simple zeros \( \lambda_1(\varepsilon), \ldots, \lambda_p(\varepsilon) \) of \( f_\varepsilon \) such that

\[
\lambda_i(\varepsilon) = \left( \frac{p! f(0)}{w^{(p)}(0)} \right)^{-1} \varepsilon + O(\varepsilon^2).
\]

(4.5)

where \( \varepsilon_i \) is the \( i \)-th complex root of the equation \( x^n = \varepsilon \). Note that \( \lim_{\varepsilon \to 0} \lambda_i(\varepsilon) = 0 \), \( \forall i \). Therefore, for sufficiently small \( \varepsilon \), we get \( f_\varepsilon \in \mathcal{F} \), \( \lambda_1(\varepsilon), \lambda_2(\varepsilon), \ldots, \lambda_p(\varepsilon) \in \mathcal{S}(f_\varepsilon) \) and \( \lim_{\varepsilon \to 0} R_\varepsilon(\lambda_i(\varepsilon)) = \min(a/2, L) \). Then the starting point \( y_0 = x_0 = 0 \) belongs to \( B(c, f) \).

This means that the sequence \( \{ y_i \} \), \( y_{i+1} = \varphi(y_i, \ldots, \varphi(f, y_i)) \), is well defined. Observe that \( \varphi(f, x_i) = \varphi(f, y_{i+1}) \) for \( j = 0, 1, \ldots, k - 1 \) and \( y_0 = x_0 = 0 \). Therefore \( y_{i+1} = x_i = \varphi(x_i, \varphi(f, x_i)) \) for \( i = 0, 1, \ldots, k - 1 \). From (4.1) we know that \( |y_i| = |x_i| \geq a/2 \), which means that \( x_i \) does not belong to the domain \( D_\varepsilon \). Thus \( \varphi(f, y_i) \) is not well defined (see (2.9)), which contradicts \( c(\varphi) > 0 \).

Theorem 4.1 says that there exists no globally convergent iteration without memory using a linear information operator for the class \( \mathcal{F} \) of analytic functions having only simple zeros. This negative result also holds for real problems. Let \( \mathcal{F} \) be the class of real functions \( f : D \subset R \to R \), analytic in \( D \) and having a unique simple zero. Consider real information operators, i.e., \( \varphi(f, x) \in R^n \), \( \forall \in \mathcal{F} \), \( \forall x \in D_e \).

**Theorem 4.2.** For any real linear information operator \( \varphi \) there exists no iteration \( \varphi \) from \( \Phi(\varphi) \) which is globally convergent for the class \( \mathcal{F} \).

**Proof.** Suppose that for a real linear information operator \( \varphi : \mathcal{F} \to D \), and for \( \varphi \in \Phi(\varphi) \), \( e = c(\varphi) > 0 \). Let \( f(x) = x - a \) be defined as in the proof of Theorem 4.1. Since \( x_0 = 0 \in B(c, f) \), the sequence \( (x_i, x_{i+1} = \varphi(x_i, \varphi(f, x_i)) \) tends to \( a \neq 0 \). Therefore

\[
x_i = \varphi(0, \varphi(f, 0)) \neq 0.
\]

(4.6)

Consider now a real polynomial \( w \):

\[
w(x) = \sum_{j=0}^{n} a_j x^{j-1}.
\]

(4.7)

satisfying

\[
(\varphi - w)(0) = \varphi(f, 0).
\]

(4.8)

The equation (4.8) is equivalent to the following homogeneous system of \( n \) real linear equations:

\[
\sum_{j=0}^{n} a_j L_j(x^{s-1}, x_0) = 0 \quad \text{for} \quad s = 1, 2, \ldots, n.
\]

(4.9)
Stationary Iteration Using Linear Information

It is obvious that there exists a nonzero real polynomial \( w(x) \) of the form (4.7) and satisfying (4.8). Let \( w(x) = x^p (x - c_1) \cdot \cdot \cdot (x - c_r) \) where \( 1 \leq r \leq 2n + 1, 1 \leq p \leq r, \) and \( c_i \neq 0 \) for \( i = p + 1, \ldots, r \). Owing to (4.7), \( p \) is odd. Define

\[
\Gamma = \begin{cases} \min \left\{ \frac{|x_1|}{2}, \frac{|x_{p+1}|}{2}, \ldots, \frac{|x_r|}{2} \right\} & \text{if } p < r, \\
\frac{|x_1|}{2} & \text{otherwise.}
\end{cases}
\]

and

\[
f_\epsilon(x) = \begin{cases} f(x) + \frac{1}{\epsilon} \cdot \eta(x) & \text{if } |x| < \Gamma, \\
\text{undefined} & \text{if } |x| \geq \Gamma.
\end{cases}
\]

for \( \epsilon > 0 \). For sufficiently small \( \epsilon \), the solutions of \( f_\epsilon(x) = 0 \) in the complex plane are given by (4.5). Since \( p \) is odd, only one of \( \lambda(\epsilon) \) in (4.5) is real. Thus \( f_\epsilon \) has a unique simple real zero, say \( \lambda(\epsilon) \). Since \( \lambda(\epsilon) \) tends to zero with \( \epsilon \), \( \lim_{\epsilon \to 0} R(\lambda(\epsilon)) = \min(\Gamma, L) \), and \( x_0 = 0 \in B(c, f) \) for sufficiently small \( \epsilon \). Therefore the sequence \( \{x_i\}_{i=0}^\infty \) is well defined. Observe that

\[
x_1 = \phi(0; f, 0) = \phi(0; \cdot)(f, 0).
\]

Owing to (4.6) \( x_i \in D_0 \) which means that \( x_i \) is not well defined. This contradiction ends the proof. \( \square \)

5. The General Case

In Section 4 we showed that there exists no globally convergent one-point stationary iteration without memory which uses linear information. In this section we prove the same result for multipoint iterations with or without memory.

Let \( L_1, L_2, \ldots, L_n \) be functionals defined as in Section 2. Then a multipoint linear information operator \( \mathcal{V} : D_\epsilon \subseteq H \times C \to C^n \), is defined as

\[
\mathcal{V}(f, x) = [L_1(f, x_1), L_2(f, x_2), \ldots, L_n(f, x_n)]. \quad \forall f \in H, \forall x \in D_\epsilon,
\]

where

\[
z_1 = x \quad \text{and} \quad z_i = \xi_1(z_1), L_1(f, z_1), L_2(f, z_2), \ldots, L_n(f, z_n), \quad k = 1, 2, \ldots, n - 1,
\]

for certain functions \( \xi_1, \xi_2, \ldots, \xi_n \). Note that if \( \xi_k = x \), \( k = 2, \ldots, n \), then \( \mathcal{V} \) is a one-point linear information operator as defined in Section 2.

For given integer \( m \), let \( x_0, x_1, \ldots, x_m \) be distinct approximations of a solution of \( f(x) = 0 \), \( f \in \mathcal{F} \). Suppose we construct a sequence of approximations by the formula

\[
x_{i+1} = \phi(x_i, x_{i-1}, \ldots, x_m; \cdot)(f, x_i, \cdot)(f, x_{i-1}, \cdot), \ldots, \cdot)(f, x_m, \cdot), \quad i = 0, 1, \ldots, m - 1
\]

where \( \phi : D_\epsilon \subseteq C^{m+1(n+1)} \to C \) is a functional. Then \( \phi \) is called a multipoint stationary iteration operator with memory if \( m \geq 1 \) and without memory if \( m = 0 \) using a linear information operator \( \mathcal{V} \). For brevity \( \phi \) is also called an iteration. Let \( \Phi_\epsilon \subseteq \Phi \) be the class of all such iterations.

For particular \( m \) and \( m \) we get commonly used iterations. For instance, \( m = 1 \) \( \cdot \) \( (f, x) = [f(x)] \), and

\[
\phi(x, y; \cdot)(f, x, \cdot)(f, y) = x - \frac{x - y}{f(x) - f(y)} f(x)
\]

is the secant iteration. An example of two-point stationary iteration without memory is provided by Steffensen iteration [2], which is defined as follows

\[
\cdot (f, x) = [f(z_1), f(z_2)] \quad \text{where} \quad z_1 = x \quad \text{and} \quad z_2 = z_1 - f(z_1)
\]
and

\[ q(x, k, f, j) = \frac{z_1 - z_2}{f(z_1) - f(z_2)} f(z_j). \]

An example of a multipoint method with memory can be found in Traub [3, pp. 185-186].

We extend the definition of global convergence as follows. Let \( \psi \) be a one- or multipoint linear information operator and let \( m \) be an integer. Recall that \( B(c, f) \) is defined in Section 3. For any iteration \( \psi \in \Phi_m(\psi) \), define \( c = c(\psi) \) as a number such that

(i) for any \( f \in \mathcal{F} \) and any choice of distinct starting points \( x_0, x_1, \ldots, x_m \) from \( B(c, f) \), the sequence \( (x_j) \), \( x_{j+1} = q(x_j, x_{j+1}, \ldots, x_{j-m}, N(f, x_j), \ldots, N(f, x_{j+1})) \) is well defined and \( \lim_{j \to \infty} x_j \in S(f) \),

(ii) for any \( \epsilon > 0 \) there exist \( f \in \mathcal{F} \) and distinct points \( x_0, x_1, \ldots, x_m \in B(c + \epsilon, f) \) such that either the sequence \( (x_j) \) is not well defined or \( \lim_{j \to \infty} x_j \notin S(f) \).

**Definition 5.1.** We shall say that an iteration \( \psi \) is globally convergent for the class \( \mathcal{F} \) if \( c(\psi) > 0 \).

Now let \( \Phi \) be the class of all stationary iterations with or without memory which use a linear information operator \( \psi \), i.e., \( \Phi = \cup \cup \Phi_m(\psi) \).

**Theorem 5.1.** No iteration \( \psi \) from \( \Phi \) is globally convergent for the class \( \mathcal{F} \).

**Proof.** Apply the proof of Theorem 4.1 with starting points

\[ x_j = \frac{j}{4(m + 1)}, \quad j = 0, 1, \ldots, m, \]

and with \( n \) in (4.2) replaced by \( n(\alpha + k) \).

Theorem 5.1 says that knowing only the value of a finite number of linear functionals on \( f \), it is impossible to find a globally convergent stationary iteration for the class \( \mathcal{F} \).

Therefore, if we want to solve \( f(1) = 0 \) by a stationary iteration, we have to assume that the starting points are sufficiently close to a solution. By contrast it is known that for some nonstationary information operators there exist globally convergent stationary iterations. An example is provided by Laguerre iteration for the class of all real polynomials with simple zeros. Also, for linear information operators, there exist globally convergent nonstationary iterations for the class \( \mathcal{F} \). An example may be adopted from Traub and Wozniakowski [4], where global convergence of the sequence of interpolatory iterations \( f_k \) is proved. For this case, the \( k \)th iteration requires the knowledge of \( k \) linear functionals of \( f \). This will be reported in Wasilkowski [6].

6. Final Comments

In Section 5 we showed that no stationary iteration using linear information is globally convergent for the class of analytic problems having simple zeros. The existence of globally convergent iterations depends on the class \( \mathcal{F} \) of functions whose zeros we want to approximate. For some simple classes there exist well-known globally convergent stationary iterations which use linear information. For example, if \( \mathcal{F} \) is the class of real functions \( f: R \to R \) whose first derivative is monotone, then Newton iteration is globally convergent.

For many interesting classes the existence of global convergent iterations is unknown. Even for the class \( \Pi \) of all real polynomials with simple real zeros, this problem is open. All known globally convergent iterations for the class \( \Pi \) are either nonstationary or use nonlinear information. For example, Newton iteration with a suitably chosen starting point, Bernoulli's method, and Laguerre iteration are globally convergent for \( \Pi \) (see, e.g., Ralston and Rabinowitz [1]). Either implicitly or explicitly these iterations use the degree \( k \) of the polynomial whose zero is desired.
Stationary Iteration Using Linear Information

Note that the degree $k$ is nonlinear information. Indeed, suppose there exists a function $g$ (in general nonlinear) such that $k = g(x, \psi(f, x))$ where $\psi$ is a multipoint linear information operator. For any $x_0$ there exists a polynomial $w$ of degree greater than one such that $\psi(w, x_0) = 0$. Taking $f(x) = x + (1/e)\psi(x)$, we get $\psi(f, x_0) = \psi(x, x_0)$ and $f \in \Pi$ for sufficiently small $e$. Therefore $g(x_0, \psi(x, x_0)) = g(x_0, \psi(f, x_0))$, but the degrees of the polynomials $x$ and $f(x)$ are different. This contradicts the assumption $k = g(x, \psi(f, x))$, with $\psi$ linear.

We believe that any globally convergent iteration for $\Pi$ has to be nonstationary or use nonlinear information. Therefore we propose the following conjecture.

Conjecture 6.1. There exists no globally convergent stationary iteration using linear information for the class of all real polynomials with simple real zeros.

ACKNOWLEDGMENTS. I thank J F. Traub and H. Wozniakowski for their help. Without it this paper could not have been completed.

REFERENCES
