THE COMPUTATIONAL COMPLEXITY OF ITERATIVE METHODS FOR SYSTEMS OF NONLINEAR EQUATIONS

Richard Brent

Mathematical Sciences Department
IBM Thomas J. Watson Research Center
Yorktown Heights, New York

I. INTRODUCTION

Suppose that an iterative method M generates successive approximations $\underset{\sim}{x}_0, \underset{\sim}{x}_1, \ldots$ to a solution $\underset{\sim}{x}^*$ of the system

$$\underset{\sim}{f}(\underset{\sim}{x}) = \underset{\sim}{0} \tag{1.1}$$

of n nonlinear equations in n unknowns. If w_i is the amount of work required to compute $\underset{\sim}{x}_i$ from $\underset{\sim}{x}_{i-1}$ (and other results saved from previous iterations), then we say that the <u>efficiency</u> of M (for the given $\underset{\sim}{f}$, $\underset{\sim}{x}_0$ etc.) is

$$E = \lim_{i\to\infty} \frac{1}{w_{i+1}} \log \left(\frac{\log ||\underset{\sim}{x}_{i+1} - \underset{\sim}{x}^*||}{\log ||\underset{\sim}{x}_i - \underset{\sim}{x}^*||} \right), \tag{1.2}$$

if the limit exists. If a method M' produces successive approximations $\underset{\sim}{x}'_i$ with work w'_i, then we say that M' has efficiency <u>at least</u> E if (1.2) holds for some w_i and $\underset{\sim}{x}_i$ satisfying $w'_i \le w_i$ and $||\underset{\sim}{x}'_i - \underset{\sim}{x}^*|| \le ||\underset{\sim}{x}_i - \underset{\sim}{x}^*||$.

Our aim is to compare the efficiencies of certain methods with the best possible, so we consider only methods with positive efficiency. For technical reasons, we assume that

$$0 < \liminf_{i\to\infty} w_i \le \limsup_{i\to\infty} w_i < \infty. \tag{1.3}$$

(Usually w_i is constant for sufficiently large i.) We assume that f has continuous partial derivatives of all orders and that the Jacobian matrix of f at x^* is nonsingular. We say that the efficiency of M (independent of a particular choice of f, x_0 etc.) is E if E is the supremum of numbers E' such that M has efficiency at least E' for all functions f (as above) and sufficiently good starting values.

The order of a method (for given f, x_0 etc.) is

$$\rho = \lim_{i\to\infty} \frac{\log \|x_{i+1} - x^*\|}{\log \|x_i - x^*\|}, \tag{1.4}$$

if the limit exists. The definitions of a method with order at least ρ, and with order ρ independent of a particular f, x_0 etc., are apparent.

The definition (1.2) has the following nice properties.

1. E is independent of the particular vector norm used (and similarly for ρ).

2. If ρ and $w = \lim_{i\to\infty} w_i$ exist, then $E = \frac{\log \rho}{w}$ is the logarithm of the "efficiency index" of Ostrowski (1960a). It follows from Gentleman (1971a) that any "reasonable" measure of computational efficiency is a monotonic function of E.

3. If methods M, M' have efficiencies E, E' and require $W(\varepsilon)$, $W'(\varepsilon)$ units of work to find an x_i such that $\|x_i - x^*\| < \varepsilon$, then

$$\lim_{\varepsilon\to 0+} \frac{W(\varepsilon)}{W'(\varepsilon)} = \frac{E'}{E} \tag{1.5}$$

Thus, M requires E'/E times as much work as M' to reduce $\|x_i - x^*\|$ to a small positive tolerance.

Except for a brief comment in Section 5, we restrict our attention to methods which depend on the sequential evaluation of $f(x)$ at certain points x, and the unit of computational work is one such evaluation. Thus, we neglect the possibility of evaluating derivatives of f except by finite differences, and any overhead, i.e., work other than function evaluations, is ignored (except in Section 4).

2. MULTIVARIATE POLYNOMIAL INTERPOLATION METHODS

Suppose that $m \ge 1$, $n \ge 1$, $N = \binom{n+m}{m} = \frac{(n+m)!}{m!n!}$, and initial distinct approximations $x_0, \ldots, x_{N-1}$ are given. The inverse polynomial interpolation method $I_{m,n}$ generates $x_N, x_{N+1}, \ldots$ in the following way. Suppose that, for some $i \ge N$, approximations $x_0, \ldots, x_{i-1}$ have been generated. Then

$$x_i = \begin{pmatrix} a^{(1)} \\ \cdot \\ \cdot \\ \cdot \\ a^{(n)} \end{pmatrix}, \tag{2.1}$$

where $a^{(1)}, \ldots, a^{(n)}$ are the constant terms in the multivariate polynomials

$$P_j(y) = a^{(j)} + \sum_{1\le k\le n} b_k^{(j)} y_k + \ldots + \sum_{1\le k_1\le \ldots \le k_m\le n} c^{(j)}_{k_1,\ldots,k_m} y_{k_1} \ldots y_{k_m} \tag{2.2}$$

which satisfy

$$x_p = \begin{pmatrix} P_1(f(x_p)) \\ \cdot \\ \cdot \\ \cdot \\ P_n(f(x_p)) \end{pmatrix} \tag{2.3}$$

for $i-N \le p < i$. (Solving the linear equations which give x_i requires of order N^2 operations if a rank-one updating method is used.)

Let

$$\varepsilon_i = \|x_i - x^*\|_2 . \tag{2.4}$$

It is shown in Section 6 that, if $\varepsilon_{i-1}, \ldots, \varepsilon_{i-N}$ are sufficiently small, then

$$\varepsilon_i \le \frac{c}{|\Delta_i|} \prod_{j=0}^{m} \max\{\varepsilon_{i-1}, \ldots, \varepsilon_{i-\binom{n+j}{j}}\}, \tag{2.5}$$

where c is a constant (depending on f), and Δ_i is a certain N by N determinant (depending on $f(x_{i-1}), \ldots, f(x_{i-N})$) of order unity.

For the moment assume that

$$\lim_{i \to \infty} \sup |\log|\Delta_i||^{\frac{1}{i}} < \rho_{m,n}, \tag{2.6}$$

where $\rho_{m,n}$ is the (unique) positive real root of

$$\sum_{j=0}^{m} \rho_{m,n}^{-\binom{n+j}{j}} = 1. \tag{2.7}$$

If $x_i \to x^*$, then it follows from (2.5) and (2.6) that the order of convergence is at least $\rho_{m,n}$. (The proof is similar to some given in Brent (1972a).) Also, there are functions and starting points such that the order is exactly $\rho_{m,n}$. Hence, $I_{m,n}$ has efficiency

$$E_{m,n} = \log \rho_{m,n}. \tag{2.8}$$

If $1 < \rho < \rho_{m,n}$ and (Δ_i) is a sequence of independent random variables distributed so that $\sum_{i=1}^{\infty} P(|\Delta_i| \le \exp(-\rho^i))$ is convergent, then (2.6) holds with probability one. This suggests that, in some sense, the order of $I_{m,n}$ is "nearly always" at least $\rho_{m,n}$. However, the order may be less than $\rho_{m,n}$ if (2.6) does not hold (and the method breaks down if $\Delta_i = 0$).

$\rho_{m,n}$ and $E_{m,n}$ are monotonic increasing functions of m, so the efficiency of $I_{m,n}$ is bounded above by

$$E_{\infty,n} = \log \rho_{\infty,n}, \tag{2.9}$$

where $\rho_{\infty,n}$ is the (unique) real positive root of

$$\sum_{j=0}^{\infty} \rho_{\infty,n}^{-\binom{n+j}{j}} = 1. \tag{2.10}$$

In view of the results of Winograd and Wolfe (1971a) for $n=1$, the following conjecture is highly plausible.

Conjecture. No locally convergent method based entirely on function evaluations has efficiency greater than $E_{\infty,n}$.

It is easy to see that no method can have efficiency greater than $\log 2$: apply Winograd and Wolfe's result to a system of equations of the form

$$\left.\begin{aligned} f_1(x_1) &= 0 \\ &\vdots \\ f_n(x_n) &= 0\ . \end{aligned}\right\} \tag{2.11}$$

However,

$$E_{\infty,n} \sim \frac{\log n}{n} \tag{2.12}$$

for large n, so our conjecture is much stronger than this.

Table 1 gives $E_{\infty,n}$ and $E_{1,n}/E_{\infty,n}$ for various values of n. Note that method $I_{1,n}$ has efficiency very close to $E_{\infty,n}$ if (2.6) holds. In fact,

$$1 - E_{1,n}/E_{\infty,n} = O(n^{-n/3}) \tag{2.13}$$

as $n \to \infty$.

Table 1: Various Efficiencies

n	$k_S(n)$	$E_{\infty,n}$	$\frac{E_{1,n}}{E_{\infty,n}}$	$\frac{E_S(n)}{E_{\infty,n}}$
1	1	0.6931	0.6942	0.6942
2	3	0.4382	0.8724	0.6817
3	4	0.3414	0.9440	0.7048
4	4	0.2880	0.9763	0.7161
5	5	0.2532	0.9908	0.7227
10	8	0.1691	1.0000	0.7285
20	12	0.1084	1.0000	0.7417
50	23	0.0568	1.0000	0.7672
100	38	0.0337	1.0000	0.7874

3. SPECIAL CASES

Some special cases of the above results are of interest. If $n = 1$, equation (2.10) reduces to

$$\sum_{j=0}^{\infty} \rho_{\infty,1}^{-(j+1)} = 1, \tag{3.1}$$

so $\rho_{\infty,1} = 2$ and $E_{\infty,1} = \log 2$. Thus, the result of Winograd and Wolfe (1971a) shows that the conjecture above is true for $n = 1$.

If $n = m = 1$, then (2.7) reduces to

$$\rho_{1,1}^{-1} + \rho_{1,1}^{-2} = 1, \tag{3.2}$$

so $\rho_{1,1} = (1+\sqrt{5})/2 = 1.618\ldots$, which is well known to be the order of the one-dimensional secant method (see Brent (1972a) or Ostrowski (1966a)). Rissanen (1971a) shows that, with certain restrictions, no method with the same memory can be more efficient.

If $n = 1$, $m \geq 1$, then (2.7) reduces to

$$\sum_{j=0}^{m} \rho_{m,1}^{-(j+1)} = 1, \tag{3.3}$$

and $\rho_{m,1}$ is the order of the well-known (direct or inverse) m-th degree polynomial interpolation methods; see Traub (1964a).

If $n \geq 1$, $m = 1$, then (2.7) reduces to

$$\rho_{1,n}^{-1} + \rho_{1,n}^{-(n+1)} = 1, \tag{3.4}$$

and $\rho_{1,n}$ is the order of Wolfe's secant method, provided (2.6) holds (this is much weaker than the assumption that Δ_i is bounded away from zero). See Wolfe (1959a), Barnes (1965a) and Bittner (1959a).

If $n = 2$ then (2.10) reduces to

$$\phi(\rho_{\infty,2}^{-1}) = 2, \tag{3.5}$$

where

$$\phi(x) = \sum_{j=0}^{\infty} x^{j(j+1)/2}. \tag{3.6}$$

We note that

$$\phi(x) = \prod_{j=1}^{\infty} \left(\frac{1 - x^{2j}}{1 - x^{2j-1}} \right) \tag{3.7}$$

by an identity of Gauss (see Hardy and Wright (1938a)). No generalization of (3.7) for $n > 2$ has been found.

4. PRACTICAL EFFICIENT METHODS

The methods $I_{m,n}$ of Section 2 are impractical if $N = \binom{m+n}{n}$ is large, for the overhead per function evaluation is of order N^2 operations. Also, their efficiency may be less than $E_{m,n}$ if (2.6) fails to hold. We shall briefly describe a class $\{S_k \mid k \geq 1\}$ of methods which avoid these disadvantages: the optimal method in the class has efficiency $E_S(n)$ close to $E_{\infty,n}$, and the overhead per function evaluation is of order n^2. (Since $\underset{\sim}{f}(\underset{\sim}{x})$ has n components $f_i(\underset{\sim}{x})$, each of which is a function of n variables, this is quite reasonable.)

If distinct approximations $\underset{\sim}{x}_i$ and $\underset{\sim}{x}'_i$ to a zero $\underset{\sim}{x}^*$ of $\underset{\sim}{f}(\underset{\sim}{x})$ have been found, then S_k generates approximations $\underset{\sim}{x}_{i+1}$ and $\underset{\sim}{x}'_{i+1}$ in the following way: if $\underset{\sim}{f}(\underset{\sim}{x}_i) = \underset{\sim}{0}$ then $\underset{\sim}{x}_{i+1} = \underset{\sim}{x}'_{i+1} = \underset{\sim}{x}_i$, otherwise do steps 1 to 4.

1. Let Q_i be an orthogonal matrix satisfying

$$\underset{\sim}{x}'_i = \underset{\sim}{x}_i + h_i Q_i \underset{\sim}{e}_1, \tag{4.1}$$

where $h_i = \| \underset{\sim}{x}_i - \underset{\sim}{x}'_i \|_2$ and

$$\underset{\sim}{e}_1 = (1, 0, \ldots, 0)^T.$$

2. Let A_i be the matrix whose j-th column is

$$A_i \underset{\sim}{e}_j = \frac{1}{h_i} [\underset{\sim}{f}(\underset{\sim}{x}_i + h_i Q_i \underset{\sim}{e}_j) - \underset{\sim}{f}(\underset{\sim}{x}_i)].$$

3. Let $\underset{\sim}{y}_{i,0} = \underset{\sim}{x}_i$ and

$$\underset{\sim}{y}_{i,j} = \underset{\sim}{y}_{i,j-1} - J_i^{-1} f(\underset{\sim}{y}_{i,j-1}) \tag{4.3}$$

for $j = 1, \ldots, k$, where $J_i = A_i Q_i^T$.

4. Let $\underset{\sim}{x}_{i+1} = \underset{\sim}{y}_{i,k}$ and $\underset{\sim}{x}'_{i+1} = \underset{\sim}{y}_{i,k-1}$.

It is shown in Brent (1972b) that the efficiency of S_k is

$$E_S(k,n) = \frac{\log \frac{1}{2}(k+\sqrt{k^2+4})}{n+k-1}. \tag{4.4}$$

If $k = k_S(n)$ is chosen so that $E_S(k,n)$ attains its maximum value $E_S(n)$, then

$$k_S(n) \sim n/\log n \tag{4.5}$$

and

$$E_S(n) \sim \frac{\log n}{n} \sim E_{\infty,n} \tag{4.6}$$

for large n. Table 1 gives $k_S(n)$ and $E_S(n)/E_{\infty,n}$ for various values of n. If the conjecture above is true, then the optimal method S_k is close to the best possible. In fact, we have

$$1 - \frac{E_S(n)}{E_{\infty,n}} = O\left(\frac{1}{\log n}\right) \tag{4.7}$$

as $n \to \infty$. It is an open question whether there are practical methods with efficiency lying between $E_S(n)$ and $E_{\infty,n}$.

5. METHODS WHICH USE COMPONENT EVALUATIONS

So far we have taken one evaluation of $\underset{\sim}{f}(\underset{\sim}{x}) = (f_1(\underset{\sim}{x}), \ldots, f_n(\underset{\sim}{x}))^T$ as the unit of computational work. If, instead, the evaluation of a component $f_i(\underset{\sim}{x})$ of $\underset{\sim}{f}(\underset{\sim}{x})$ is taken as $\frac{1}{n}$ units of work, then methods with efficiency greater than $E_{\infty,n}$ exist (at least for $n \geq 10$). In Brent (1972b) we describe a class $\{T_k \mid k \geq 1\}$ of methods related to Brown's method (see Brown and Conte (1967a), Rabin (1972a)). The optimal method in this class has efficiency

$$E_T(n) = \max_{k=1,2,\ldots} \frac{2\log(k+1)}{n+2k+1}, \tag{5.1}$$

and

$$E_T(n) \sim 2E_{\infty,n} \tag{5.2}$$

for large n. Whether significantly more efficient methods exist is an open question.

6. APPENDIX

In this appendix we sketch a proof of the inequality (2.5). Let $\underset{\sim}{g}$ be the inverse function of $\underset{\sim}{f}$, so

$$\underset{\sim}{g}(\underset{\sim}{f}(\underset{\sim}{x})) = \underset{\sim}{x} \tag{6.1}$$

for all $\underset{\sim}{x}$ sufficiently close to the simple zero $\underset{\sim}{x}^*$ of $\underset{\sim}{f}$. Let

$$\underset{\sim}{y}^{(p)} = \underset{\sim}{f}(\underset{\sim}{x}_p) \tag{6.2}$$

and

$$\eta_p = \|\underset{\sim}{y}^{(p)}\|_2 \tag{6.3}$$

for $p = i-1, \ldots, i-N$. By a renumbering, if necessary, there is no loss of generality in assuming that

$$\eta_{i-1} \leq \eta_{i-2} \leq \ldots \leq \eta_{i-N}. \tag{6.4}$$

Let g_j be the j-th component of $\underset{\sim}{g}$, and $y_k^{(p)}$ the k-th component of $\underset{\sim}{y}^{(p)}$. By equations (2.2), (2.3), (6.1) and (6.2),

$$g_j(\underset{\sim}{y}^{(p)}) = a^{(j)} + \sum_{1 \leq k \leq n} b_k^{(j)} y_k^{(p)} + \ldots + \sum_{1 \leq k_1 \leq \ldots \leq k_m \leq n} c^{(j)}_{k_1,\ldots,k_m} y^{(p)}_{k_1} \ldots y^{(p)}_{k_m} \tag{6.5}$$

for $1 \leq j \leq n$ and $i-N \leq p < i$.

Compare (6.5) with the Taylor series expansion

$$g_j(y) = A^{(j)} + \sum_{1\le k\le n} B_k^{(j)} y_k + \cdots + \sum_{1\le k_1\le\ldots\le k_m\le n} C^{(j)}_{k_1,\ldots,k_m} y_{k_1}\cdots y_{k_m} + R_j(y) \qquad (6.6)$$

of g_j about 0. If $a^{(j)} = a^{(j)} - A^{(j)}$ etc., then putting $y = y^{(p)}$ in (6.6) and subtracting (6.5) gives

$$a^{(j)} + \sum_{1\le k\le n} \beta_k^{(j)} y_k^{(p)} + \ldots + \sum_{1\le k_1\le\ldots\le k_m\le n} \gamma^{(j)}_{k_1,\ldots,k_m} y^{(p)}_{k_1}\cdots y^{(p)}_{k_m} = R_j(y^{(p)}) \qquad (6.7)$$

for $1\le j\le n$ and $i-N\le p< i$. For each j, this gives a system of N linear equations in the N variables $a^{(j)}, \beta_1^{(j)}, \ldots, \gamma^{(j)}_{n,\ldots,n}$. Solving by Cramer's rule for $a^{(j)}$ gives

$$a^{(j)} = D_1^{(j)} / D_2, \qquad (6.8)$$

where $D_1^{(j)}$ and D_2 are N by N determinants.

From the assumption (6.4) and the observation that $R_j(y^{(p)})$ is of order η_p^{m+1}, an inspection of the dominant terms in (6.8) shows that

$$|a^{(j)}| \le \frac{K_j}{|\Delta_i|} \prod_{k=0}^{m} \eta_{i-(n+k \atop k)}, \qquad (6.9)$$

where K_j is a constant, and

$$\Delta_i = D_2 \prod_{k=1}^{m} \prod_{j=1+\binom{n+k-1}{k-1}}^{\binom{n+k}{k}} \eta_{i-j}^{-k} \qquad (6.10)$$

is of order unity.

From (6.1) and (6.6), it is immediate that the zero x^* of f is given by

$$x^* = \begin{pmatrix} A^{(1)} \\ \vdots \\ A^{(n)} \end{pmatrix}. \qquad (6.11)$$

Thus, from (2.1) and (6.9), we have

$$\|x_i - x^*\|_2 \le \frac{K}{|\Delta_i|} \prod_{k=0}^{m} \eta_{i-(n+k \atop k)}, \qquad (6.12)$$

where

$$K = \left(\sum_{j=1}^{n} K_j^2\right)^{\frac{1}{2}}. \qquad (6.13)$$

In view of the assumption (6.4) and the fact that $x_{i-1}, \ldots, x_{i-N}$ are close to the simple zero x^*, the result (2.5) follows from (6.12).

References

Barnes, J. G. P., 1965a, An algorithm for solving non-linear equations based on the secant method, Comp. J. 8, 66-72.

Bittner, L., 1959a, Eine Verallgemeinerung des Sekantenverfahrens (regula falsi) zur naherungsweisen Berechnung der Nullstellen eines nichtlinearen Gleichungssystems, Wissen. Zeit. der Technischen Hochschule Dresden, Vol. 9, 325-329.

Brent, R. P., 1972a, Algorithms for Minimization without Derivatives, Prentice-Hall, Englewood Cliffs, New Jersey (Chapter 3).

Brent, R. P., 1972b, On maximizing the efficiency of algorithms for solving systems of nonlinear equations, IBM Watson Research Center, Yorktown Heights, New York, RC 3725.

Brown, K. M. and Conte, S. D., 1967a, The solution of simultaneous nonlinear equations, Proceedings 22nd National Conference of the ACM, Thompson Book Co., Washington, D. C., 111-114.

Gentleman, M. W., 1971a, Private communication.

Hardy, G. H. and Wright, E. M., 1938a, An Introduction to the Theory of Numbers, Clarendon Press, Oxford, Thm. 354.

Ostrowski, A. M., 1960a, Solution of Equations and Systems of Equations, Academic Press, New York.

Rabin, M. O., 1972a, Solving linear equations by means of scalar products, these Proceedings, 11-20.

Rissanen, J., 1971a, On optimum root-finding algorithms, J. Math. Anal. Appl. 36, 220-225.

Winograd, S. and Wolfe, P., 1971a, Optimal iterative processes, IBM Watson Research Center, Yorktown Heights, New York, RC 3511.

Wolfe, P., 1959a, The secant method for simultaneous nonlinear equations, Comm. ACM 2, 12-13.