数理解析研究所講究録
915巻 1995年 218-224

An Experimental Study of the Starting Values of the Durand-Kerner-Aberth Iteration

Kazufumi OZAWA (小澤一文)
Education Center for Information Processing
Tohoku University
E-mail: ozawa@dais.is.tohoku.ac.jp

1 Introduction

When solving non-linear equations by iterative methods the number of iterations depends greatly on the starting values. In this note, we will present some simple procedure to determine the starting values which reduces the number of iterations of the Durand-Kerner-Aberth method for solving algebraic equations.

2 Convergence of Durand-Kerner iteration

Let us consider the algebraic equation

$$p(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n = a_0(z-\alpha_1)\cdots(z-\alpha_n) = 0. \tag{1}$$

The Durand-Kerner-Aberth (in short D-K-A) iteration [1] for solving the equation (1) is given by

$$z_i^* = z_i + \Delta_i, \quad i = 1, \cdots, n, \tag{2}$$

where

$$\Delta_i = -\frac{p(z_i)/p'(z_i)}{1 - p(z_i)/p'(z_i)\sum\limits_{\substack{j=1 \\ j\neq i}}^{n}(z_i - z_j)^{-1}}. \tag{3}$$

The D-K-A iteration is known to converge with a cubical order of convergence when the starting values are sufficiently close to the solutions. To derive the procedure for obtaining "good" starting values of the D-K-A method we investigate the behaviors of z_i's before they converge with the cubical order of convergence.

We now suppose that after the several iterations many of the approximate solutions z_i $(i = 1, \ldots, n)$ are in small errors but the remainders are in large errors. To be more specific, suppose that the set of the indices $N = \{1, \ldots, n\}$ can be divided into the two sets N_0 and N_1 according to the sizes of the errors, i.e., that for the two sets of indices

$$N_0 = \{1, \cdots, n-m\}, \quad N_1 = \{n-m+1, \cdots, n\}, \quad N = N_0 \cup N_1,$$

where $0 < m \ll n$, the errors $e_i = z_i - \alpha_i$ $(i = 1, \ldots, n)$ of the approximate solutions satisfy the condition

$$e_i = \begin{cases} O(\rho_0), & i \in N_0, \\ O(\rho_1), & i \in N_1, \end{cases} \tag{4}$$

where ρ_0 and ρ_1 are the positive constants such that

$$0 < \rho_0 \ll \rho_1, \quad 0 < \rho_0 \ll 1.$$

Suppose, moreover, that the minimum distance of the approximate solutions in each group is relatively large compared with ρ_0 and ρ_1, i.e., that the distances D_0 and D_1 defined by

$$D_0 = \min_{\substack{i,j\in N_0 \\ i\neq j}} |z_i - z_j|, \quad D_1 = \min_{\substack{i,j\in N_1 \\ i\neq j}} |z_i - z_j|$$

satisfy that

$$0 < c_0 = \rho_0 D_0^{-1} < 1, \quad 0 < c_1 = \rho_1 D_1^{-1} < 1.$$

Under these assumptions we have the following result for the errors e_i^* of the next step[3]:

$$\begin{array}{ll} e_i^* = O(\rho_0^2), & i \in N_0, \\ |e_i^*| \leq (m-1)\ c_1^2\,(1+O(c_1))\,|e_i| + O(\rho_0), & i \in N_1. \end{array} \tag{5}$$

We can conclude that:

- Each of the approximate solutions belonging to the group N_0 converges quadratically.
- On the other hand, if $(m-1)c_1^2 < 1$ then the rate of convergence of each member of the group N_1 is not quadratic but linear. In particular, if $m = 1$ then the error of this group reduces up to the same order as that of the group N_0.

The result above suggests that it is more advantageous to start the D-K-A iteration with the set of starting values such that the number of starting values with small errors are large. Consequently, if we specify the starting values on the circle centered at the average of the roots then it is better to take the circle with a moderate radius.

Thus, the author recommends the following set of the starting values:

$$z_i^{(0)} = \beta + r_{gm} \exp\left[\sqrt{-1}\left(\frac{2\pi(i-1)}{n} + \frac{3}{2n}\right)\right], \quad i = 1, \cdots, n, \tag{6}$$

where β is the arithmetic mean of the roots and r_{gm} is the geometric mean of the deviations $|\alpha_i - \beta|$ $(i = 1, \ldots, n)$:

$$\beta = -\frac{a_1}{n a_0}, \quad r_{gm} = \left|\frac{p(\beta)}{a_0}\right|^{1/n}.$$

Note that the set of the starting values (6) is much more easily calculable than is the Aberth's one.

3 Numerical experiments

3.1 Equation with random coefficients

In order to show the efficiency of our starting values, we solve 50 equations of degree 40 of the following types:

- The equation with random coefficients in the complex region $(-5, 5)\times(-5, 5)$, keeping the leading coefficient $a_0 = 1$.
- The equation with the random roots in the complex region $(-1, 1) \times (-1, 1)$.

For each of the equations above, we compare the cpu-time of the D-K-A method with that of the D-K(2nd order Durand-Kerner)[2] method and those of the other codes. This experiment is performed on an ACOS 3900 computer using the double precision arithmetic of FORTRAN 77. The results are shown in Table 1, 2.

Table 1. Comparison of cpu-times(sec) for the equations with random coefficients.

	mean	var(sec^2)	max	min
DKA(A)	1.3535E-01	4.6503E-04	1.7749E-01	9.4884E-02
DKA(G)	3.1951E-02	3.8267E-06	3.5896E-02	2.7601E-02
DK(A)	1.9623E-01	1.1468E-03	2.9514E-01	1.2931E-01
DK(G)	1.0323E-01	1.0335E-03	1.8930E-01	5.2254E-02
DZPOCC	3.9729E-02	4.2467E-06	4.7141E-02	3.5502E-02
POLEQB	2.4044E-02	2.0510E-06	2.6911E-02	2.0076E-02

Table 2. Comparison of cpu-times(sec) for the equations with random roots.

	mean	var(sec^2)	max	min
DKA(A)	1.0198E-01	1.7735E-04	1.4346E-01	8.1446E-02
DKA(G)	4.1480E-02	8.8028E-06	4.8637E-02	3.6441E-02
DK(A)	1.4647E-01	3.9726E-04	2.0425E-01	1.1525E-01
DK(G)	5.4503E-01	1.0053E-01	1.5402E+00	2.6291E-01
DZPOCC	3.9549E-02	4.2655E-06	4.5209E-02	3.5883E-02
POLEQB	1.9913E-02	6.6867E-07	2.1444E-02	1.8431E-02

DKA(A): Durand-Kerner-Aberth method started with the Aberth's initial value.
DKA(G): Durand-Kerner-Aberth method started with the initial values on the circle centered with β radius r_{gm}.
DK(A): Durand-Kerner method started with the Aberth's initial value.
DK(G): Durand-Kerner method started with the initial values on the circle centered with β radius r_{gm}.
DZPOCC: IMSL library.
POLEQB: NUMPAC library.

We can see that our procedure for the starting values works well for the Durand-Kerner-Aberth iteration, but not so well for the Durand-Kerner iteration.

3.2 Costs of iteration

To measure the costs of the iterations we consider the two functions defined by

$$\mathrm{CAM} = N_{am} \sum_{i=1}^{n} (r_i - r)^2, \tag{7}$$

$$\mathrm{CGM} = N_{gm} \sum_{i=1}^{n} (\log r_i - \log r)^2, \tag{8}$$

$$r_i = |\alpha_i - \beta|, \quad i = 1, \ldots, n.$$

In Eqs. (7) and (8) the constants N_{am} and N_{gm} are chosen so that each of the functions takes 1 as its minimum. One can easily see that the values of r at which these two cost functions take their minimum are given by r_{am} and r_{gm}, the arithmetic mean and the geometric mean of r_i's, respectively.

In what is to follow, we solve the equations with given roots starting with the initial values on various circles, and calculate the values of the cost functions and the number of iterations for each of the circles.

3.2.1 Equation with uniformly distributed roots

In this experiment we solve the equation of degree 40 with uniformly distributed random roots in $(0,1) \times (0,1)$. The radii of the equation are $r_{ab} = 0.888$, $r_{am} = 0.388$ and $r_{gm} = 0.342$.

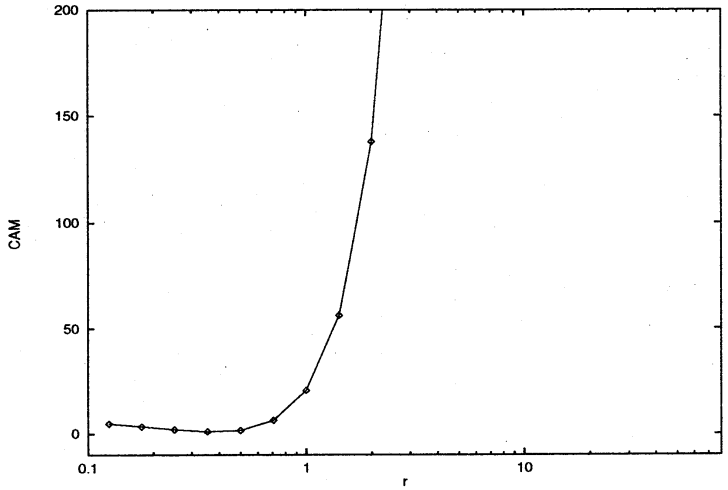

Fig. 1.1. Cost function CAM versus radius r.

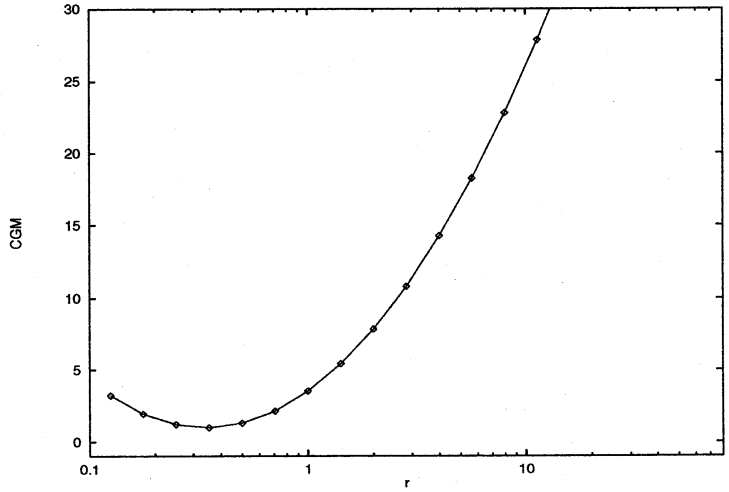

Fig. 1.2. Cost function CGM versus radius r.

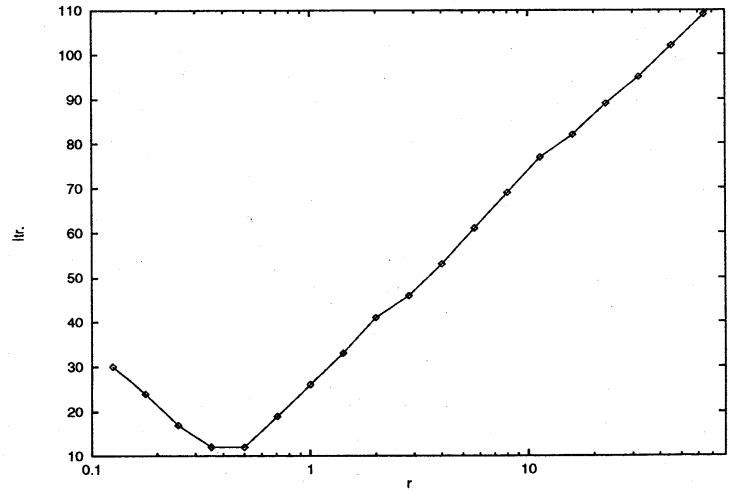

Fig. 1.3. Number of iterations versus radius r.

3.2.2 Equation with uniformly distributed roots and with several large roots

In this experiment, as before, we solve the equation of degree 40 with uniformly distributed random roots in $(0,1) \times (0,1)$, but two of these are replaced with ± 100. The radii of the equation are $r_{ab} = 100$, $r_{am} = 10.4$ and $r_{gm} = 0.461$.

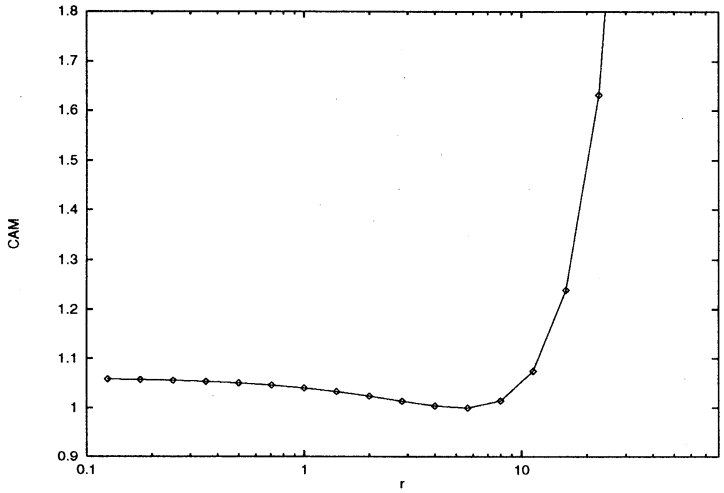

Fig. 2.1. Cost function CAM versus radius r.

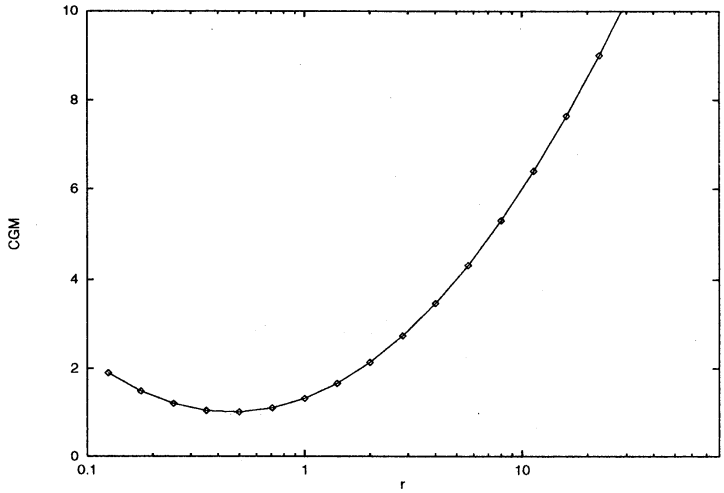

Fig. 2.2. Cost function CGM versus radius r.

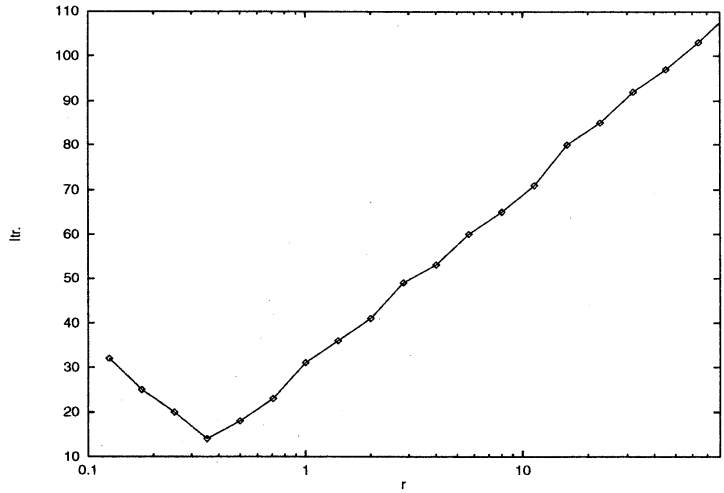

Fig. 2.3. Number of iterations versus radius r.

3.2.3 Equation with real roots

The equation solved in this experiment is the Legendre polynomial of degree 40. The radii of the equation are $r_{ab} = 3.73$, $r_{am} = 0.635$ and $r_{gm} = 0.949$.

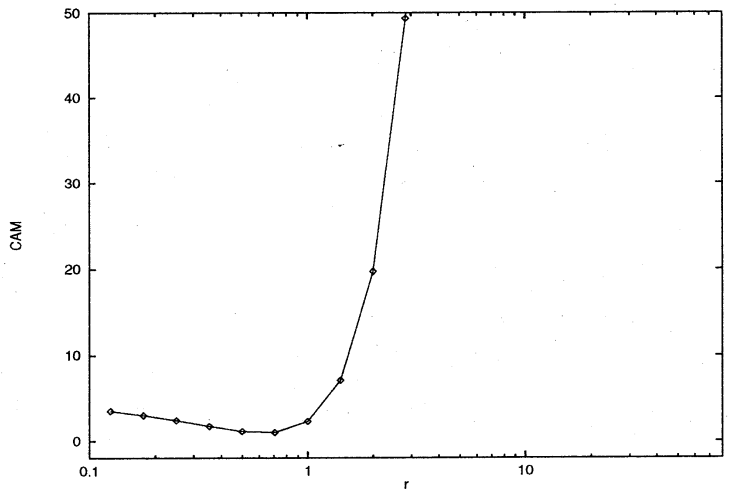

Fig. 3.1. Cost function CAM versus radius r.

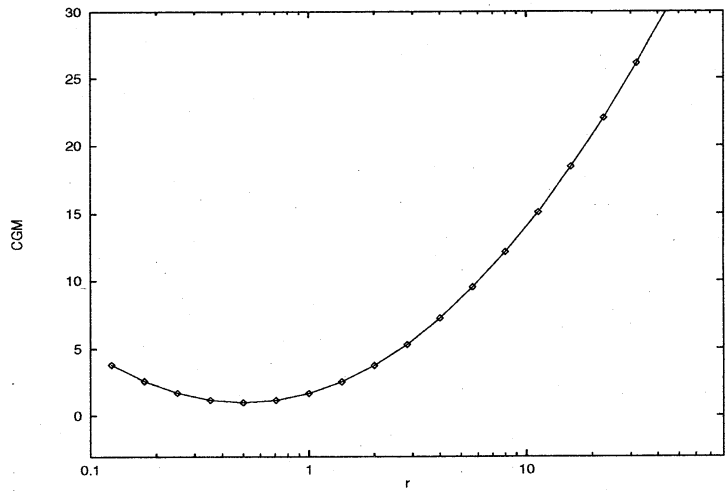

Fig. 3.2. Cost function CGM versus radius r.

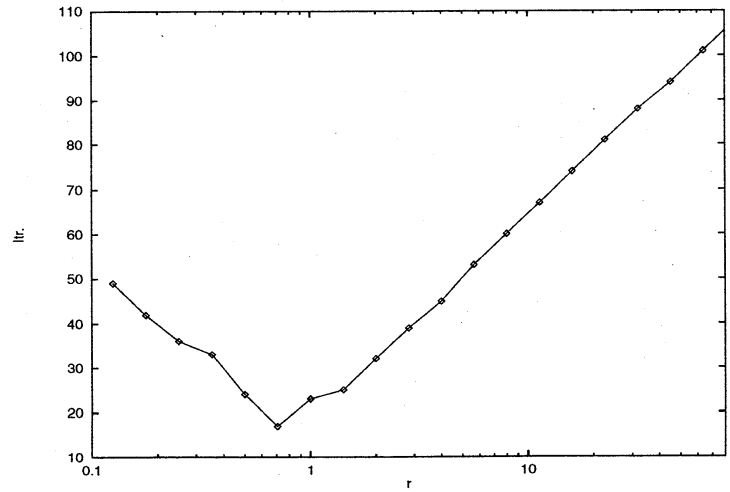

Fig. 3.3. Number of iterations versus radius r.

From the experiments above, we can see that the geometric mean r_{gm} approximates well the radius r which minimizes the iteration for the various type of equations. On the other hand, the arithmetic mean r_{am} and the Aberth's radius r_{ab} does not necessary approximate this radius. Therefore our procedure using geometric mean works well for various types of the equations. A more detailed discussion is found in [3].

References

[1] Aberth, O.,"Iteration Methods for Finding all Zeros of a Polynomial Simultaneously, "Numer. Math. **27**(1973),339-344.

[2] Kerner, I. O., "Ein Gesamtschrittverfahren zur Berechung der Nullstellen von Polynome," Numer. Math.**8** (1960), 290-294.

[3] Ozawa, K.,"A Simple Procedure for Setting the Efficient Starting Values of the Durand-Kerner Iteration(in Japanese)," Transaction of the JSIAM **3**(1993),173-186.