On the Aitken's δ^2-process

Ken Iguchi*

Abstract

This paper is concerned with the Aitken's δ^2-process and similar process. The Aitken's δ^2-process is not always successful for the sequence converging linearly.

We develop the general theory of the Aitken's δ^2-process, and we derive effective process even when the Aitken's δ^2-process fails. As an example of these processes, we consider an application to finding the absolutely largest eigenvalue by the power method. Some test matrices and its results are given.

1. Introduction

Let $\{x_p\}$ be a real sequence which converges to a limit x, where a suffix in x_p takes positive integer. Then x_p denote the p-th value of the sequence.

Suppose that

$$x_p = x + c_1 t_1^p + c_2 t_2^p + O(t_3^p), \tag{1.1}$$

where c_1 and c_2 are constants, and real numbers t_1, t_2, and t_3 are supposed to be

$$1 > |t_1| > |t_2| \geqq |t_3|. \tag{1.2}$$

Here we propose an accelerating process which is defined by

$$x_{p+2}^{(n)} = x_{p+2} + \omega_n (x_{p+2} - x_p), \tag{1.3}$$

in terms of successive terms x_p, x_{p+1}, and x_{p+2}, where

$$\omega_n = \sum_{i=1}^{n} t^{2i}, \ n = 1, 2, \ldots, \tag{1.4}$$

$$t = (x_{p+2} - x_{p+1})/(x_{p+1} - x_p). \tag{1.5}$$

The value $x_{p+2}^{(n)}$ is considered to be an approximation to the limit x.

In practice, this process is used as follows; an iterative procedure with an initial vector x_0 give successive iterates x_1, x_2, and x_3 which are used to obtain a new initial value $x_2^{(n)}$ by means of eq. (1.3). Replacing $x_2^{(n)}$ by x_0, that is, $x_0 = x_2^{(n)}$, we deduce x_1', x_2', and x_3' from x_0 by the iterative procedure. Replacing x_1', x_2', and x_3' by x_1, x_2, and x_3, that is, $x_1 = x_1'$, $x_2 = x_2'$, and $x_3 = x_3'$, use eq. (1.3) to obtain $x_2^{(n)}$, and so on.

Now taking the limit as $n \to \infty$ of eq. (1.4), we obtain the formula as

$$\omega_\infty = t^2/(1 - t^2). \tag{1.6}$$

This case of $n \to \infty$ of eq. (1.3) is usually called the Aitken's δ^2- process(see [1],

This paper first appeared in Japanese in Joho–Shori (Journal of the Information Processing Society of Japan), Vol. 15, No. 11 (1974), pp. 836~840.

* Toyota Technical College

[4], [5], [6], [8], etc.).

In the case of $|t_2/t_1| \approx 1$, the employment of the Aitken's δ^2-process makes convergence worse, and, in the worst, causes an oscillation far from the limit x. However, the employment of eq.(1.3) for low number of n, say, n = 1, 2 results in the rapid convergence when the Aitken's δ^2-process is failed.

Here we propose to use these accelerating processes as follows;

(i) if $|t_2/t_1| < 0.4$, we use the formula as

$$x_{p+2}^{(\infty)} = x_{p+2} + \omega_\infty(x_{p+2} - x_p), \tag{1.7}$$

(ii) if $0.4 \leqq |t_2/t_1| < 0.9$, we use the formula as

$$x_{p+2}^{(2)} = x_{p+2} + \omega_2(x_{p+2} - x_p), \tag{1.8}$$

(iii) if $0.9 \leqq |t_2/t_1|$, we use the formula as

$$x_{p+2}^{(1)} = x_{p+2} + \omega_1(x_{p+2} - x_p), \tag{1.9}$$

where

$$\omega_1 = t^2, \quad \omega_2 = t^2 + t^4, \tag{1.10}$$

and t is given by eq.(1.5).

The ratio of t_2/t_1 is estimated in the following way.

We evaluate t_k^*, k = 1, 2, 3 from successive five terms as follows,

$$t_k^* = (x_{p+k+1} - x_{p+k})/(x_{p+k} - x_{p+k-1}), \; k = 1,2,3. \tag{1.11}$$

Then the ratio of t_2/t_1 is estimated by means of the formula as

$$t_2/t_1 \approx (t_3^* - t_2^*)/(t_2^* - t_1^*). \tag{1.12}$$

It is sufficient to estimate only once the magnitude of the ratio of t_2/t_1 before starting calculations.

Since eq.(1.3) is based on eq.(1.1), it is applicable to the power method for the largest eigenvalue in magnitude, and the iterative methods, say, S.O.R method for solution of a system of linear equations. In the later section, we consider the power method as a typical example.

2. Accelerating process

We derive an accelerating process defined by eq.(1.3).

From eq.(1.1), we obtain the formula as

$$x_{P+1} - x_p = ct_1^p + \varepsilon_p, \tag{2.1}$$

where

$$c = c_1(t_1 - 1), \qquad \varepsilon_p = c_2t_2^p(t_2 - 1) + O(t_3^p). \tag{2.2}$$

It follows that ε_p tends asymptotically to zero as p tends to infinity, since $|t_2| < 1$.

The eq.(2.1) can be rewritten as follows;

Lemma 1

$$x_{p+1} = x_p + ct^p, \tag{2.3}$$

for any positive integer p, where c is constant, and t tends asymptotically to t_1 as p tends to infinity.

The repeated application of Lemma 1 yields the following Lemma 2.

Lemma 2

$$x_{p+m} = x_p + ct^p\left(\sum_{i=0}^{m-1} t^i\right), \tag{2.4}$$

for any positive integer p and m.

The application of Lemma 1 and Lemma 2 yields the following Lemma 3.

Lemma 3

$$x_{p+m} = x_{p+m-1} + ct^{p+m-1}, \tag{2.5}$$

for any positive integer p and m.

The eq.(2.5) corresponds to eq.(2.3) which is the case of $m = 1$ in eq.(2.5).

The repeated application of Lemma 3 yields the following Lemma 4.

Lemma 4

$$x_{p+n} = x_{p+m} + ct^{p+m}\left(\sum_{i=0}^{n-m+1} t^i\right), \tag{2.6}$$

for any positive integer p, m, and n.

From eq.(2.4) and eq.(2.5), we obtain the formula as

$$\gamma = (x_{p+m} - x_p)/(x_{p+m} - x_{p+m-1}) \tag{2.7}$$

$$= \left(\sum_{i=0}^{m-1} t^i\right)/t^{m-1}. \tag{2.8}$$

The eq.(2.8) can be rewritten in the form as

$$(1 - \gamma)t^{m-1} + \sum_{i=0}^{m-2} t^i = 0, \tag{2.9}$$

where t is given by eq.(2.7).

By eliminating c in eq.(2.4) and eq.(2.6), we obtain the following theorem.

Theorem

$$x_{p+n} = x_{p+m} + (t^m(1 - t^{n-m})/(1 - t^m))(x_{p+m} - x_p), \tag{2.10}$$

for any positive integer p, m, and n ($n > m \geqq 2$), where t is the unique real root with magnitude smaller than unity of the equation of eq.(2.9) in term of γ given by eq.(2.7).

Taking the limit as $n \to \infty$ of eq.(2.10) with $m = 2$, and then putting $x_{p+2}^{(\infty)} = \lim_{n\to\infty} x_{p+n}$, we obtain the following Corollary 1.

Corollary 1

$$x_{p+2}^{(\infty)} = x_{p+2} + (t^2/(1 - t^2))(x_{p+2} - x_p), \tag{2.11}$$

for any positive integer p, where t is given by

$$t = (x_{p+2} - x_{p+1})/(x_{p+1} - x_p). \tag{2.12}$$

Taking $m = 2$, $n = 2k$ ($k = 1,2,\ldots$) of eq.(2.10) and then putting $x_{p+2}^{(k)} = x_{p+2k}$, we obtain the following Corollary 2.

Corollary 2

$$x_{p+2}^{(k)} = x_{p+2} + \left(\sum_{i=1}^{k-1} t^{2i}\right)(x_{p+2} - x_p), \tag{2.13}$$

for any positive integer p and $k(\geqq 2)$, where t is given by eq.(2.12).

3. Estimation of the ratio of t_2/t_1

We derive a formula for estimating the magnitude of the ratio of t_2/t_1.

From successive five terms x_p, x_{p+1}, x_{p+2}, x_{p+3}, and x_{p+4}, we evaluate t_k^*, $k = 1,2,3$ by means of eq.(1.11).

By substituting eq.(1.1) into eq.(1.11), we obtain the formula as

$$t_k^* = (At_1^{p+k} + Bt_2^{p+k})/(At_1^{p+k-1} + Bt_2^{p+k-1}),\ k = 1,2,3, \tag{3.1}$$

where

$$A = c_1(t_1 - 1), \quad B = c_2(t_2 - 1). \tag{3.2}$$

By evaluating approximately eq.(3.1), we obtain the formula as

$$t_k^* \simeq t_1(1 + (B/A)(t_2/t_1)^{p+k})(1 - (B/A)(t_2/t_1)^{p+k-1})$$

$$\simeq t_1(1 - (B/A)(t_2/t_1)^{p+k-1}),\ k = 1,2,3. \tag{3.3}$$

From eq.(3.3), we obtain the formula as

$$t_{k+1}^* - t_k^* \simeq (B/A)(t_2/t_1)^{p+k-1} - (B/A)(t_2/t_1)^{p+k},\ k = 1,2. \tag{3.4}$$

Then we obtain the eq.(1.12) from eq.(3.4).

4. Application to the power method

As a typical example of the use of the accelerating process described in the preceding section, we consider an application to the power method for the largest eigenvalue in magnitude. Then it is necessary to apply these processes to all the components of the eigenvector. This can be done by the following way(see[2], [6], and [8], etc.).

Here we assume that the n x n matrix A is symmetrical and real with the eigenvalues , λ_i, i = 1,2,...,n.

Further assume that

$$|\lambda_1| > |\lambda_2| > |\lambda_3| \geqq |\lambda_4| \geqq \cdots \geqq |\lambda_n|. \tag{4.1}$$

We take an initial vector, u_0 arbitraryly.

By successive premultiplication of the matrix A form the sequence

$$u_1 = Au_0, \quad u_2 = A^2u_0, \quad u_3 = A^3u_0. \tag{4.2}$$

Then the Rayleigh quotient is

$$\bar{\lambda}_1 = (A^3u_0, A^3u_0)/(A^2u_0, A^3u_0). \tag{4.3}$$

Dividing u_1, u_2, and u_3 of eq.(4.2) by $\bar{\lambda}_1$, $\bar{\lambda}_1^2$, and $\bar{\lambda}_1^3$ respectively, we again denote these resulting vectors by u_1, u_2, and u_3. Then we apply the accelerating process of eq.(1.3) to the iterated vectors, u_1, u_2, and u_3 as

$$u = u_3 + \omega_n(u_3 - u_1), \tag{4.4}$$

where, as the value of ω_n, we take the one value among ω_1, ω_2, and ω_∞ given by eq.(1.10) and eq.(1.6).

Here we calculate the value of t for the only one component as follows;

$$t = (u_3^{(r)} - u_2^{(r)})/(u_2^{(r)} - u_1^{(r)}), \tag{4.5}$$

where r denotes the r-th component of the vectors, u_i, i = 1,2,3 which is the largest component in magnitude.

Finally replacing u_0 by u, that is, $u_0 = u$, this procedure is repeated until the desired accuracy is obtained.

5. Test matrices and numerical results

The following matrices demonstrate the effectiveness of the accelerating processes described in the section 2 according to the procedure of the section 3. The computations were carried out on the FACOM 230-60 computer of the University of Nagoya, with the single precisions of the significant figures of 7.8 in decimal.

We take $u_0^T = (1,1,\ldots,1)$ as an initial vector for all the test matrices.

Example 1[3]

Let

$$A_1 = \begin{pmatrix} 10 & 1 & 2 & 3 & 4 \\ 1 & 9 & -1 & 2 & -3 \\ 2 & -1 & 7 & 3 & -5 \\ 3 & 2 & 3 & 12 & -1 \\ 4 & -3 & -5 & -1 & 15 \end{pmatrix}$$

Then $\lambda_1 \simeq 19.1754203$,

$t_1(=\lambda_2/\lambda_1) \simeq 0.82$, $t_2/t_1(=\lambda_3/\lambda_2) \simeq 0.59$.

Example 2[7]

Let

$$A_2 = \begin{pmatrix} 0.45 & 0.55 & 0.40 \\ 0.55 & 0.45 & -0.40 \\ 0.40 & -0.40 & 0.50 \end{pmatrix}$$

Then $\lambda_1 = 1$,

$t_1 = -0.9$, $t_2/t_1 \simeq -0.33$.

We denote here the number of the iterations by N.I., and the number of the use of the accelerating processes by N.A..

Table. 1 The absolutely largest eigenvalue of the matrix A_1: $\lambda_1=19.1754203$

N.I.	N.A.	O	ε^2	$\varepsilon^2+\varepsilon^4$	$\varepsilon^2/(1-\varepsilon^2)$
3	0	15.83796	15.83796	15.83796	15.83796
6	2	15.96399	18.38206	18.36523	15.87066
9	9	16.23291	19.15411	19.16870	16.02286
12	3	16.86787	19.17257	19.17508	16.41739
15	4	17.80773	19.17492	19.17503	18.06301
18	5	18.58001	19.17535	19.17538	17.46909
21	6	18.96264	19.17540	19.17541	16.42896
24	7	19.10557	19.17541	19.17542	18.10049
27	8	19.15317	19.17541		17.55771
30	9	19.16840	19.17542		16.54619

Table. 2 The absolutely largest eigenvalue of the matrix A_2: $\lambda_1=1.0$

N.I.	N.A.	O	ε^2	$\varepsilon^2+\varepsilon^4$	$\varepsilon^2/(1-\varepsilon^2)$
3	0	1.465038	1.465038	1.465038	1.465038
6	1	1.221964	1.220619	1.220612	1.220612
9	2	1.111839	1.106536	1.106268	1.106253
12	3	1.057840	1.050494	1.049690	1.049570
15	4	1.030306	1.022933	1.021631	1.021249
18	5	1.015986	1.009680	1.008476	1.007791
21	6	1.008462	1.004036	1.002926	1.002157
24	7	1.004488	1.001573	1.000688	1.000350
27	8	1.002382	1.000591	1.000242	1.000018
30	9	1.001265	1.000216	1.000061	1.000000

Aknowledgement

The author would like to thank the referee for his helpful remarks for this paper.

References

1) A. C. Aitken; On Bernoulli's numerical solution of algebraic equations. Proc. Royal Soc. Edinburgh 46, 289-305 (1926).

2) V. N. Faddeeva; Computational Method of Linear Algebra. Dover Publications, New York, 1959.

3) R. T. Gregory & D. L. Karney; A Collection of Matrices for Testing Computational Algorithms. Wiley-Interscience, New York, 1969.

4) S. Hitotumatu; Numerical Analysis. Zeimu-Keiri-Kyokai, 1971.

5) B. Noble; Numerical Method:1. Oliver & Boyd, Edinburgh, 1964.

6) A. Ralston; A First Course in Numerical Analysis. McGraw-Hill, New York, 1965.

7) J. W. Schmidt; Asymptotische Einschliessung bei Konvergenzbeschleeunigenden Verfahren. Num. Math., 8, 105-113 (1966).

8) J. F. Traub; Iterative Methods for the Solution of Equations. Prentice-Hall, Englewood Cliffs, N. J., 1964.

9) J. H. Wilkinson; The Algebraic Eigenvalue Problem. Oxford University Press, London, 1965.