OK

NASA Technical Memorandum 82931

Acceleration of Convergence of Vector Sequences

(NASA-TM-82931) ACCELERATION OF CONVERGENCE OF VECTOR SEQUENCES (NASA) 10 p
HC A02/MF A01 CSCL 12A
N82-29075
Unclas
G3/64 28375

Avram Sidi and William F. Ford
Lewis Research Center
Cleveland, Ohio

and

David A. Smith
Duke University
Durham, North Carolina

Prepared for the
Thirtieth Anniversary Meeting of the Society
for Industrial and Applied Mathematics
Stanford, California, July 19-23, 1982

NASA

ACCELERATION OF CONVERGENCE OF VECTOR SEQUENCES

Avram Sidi* and William F. Ford
National Aeronautics and Space Administration
Lewis Research Center
Cleveland, Ohio 44135

and

David A. Smith
Duke University
Durham, North Carolina 27706

ABSTRACT

A general approach to the construction of acceleration of convergence methods for vector sequences is proposed. Using this approach one can generate some known methods and some new ones. In this talk we shall concentrate on one new method, which turns out to be a simplified version of minimal polynomial extrapolation. We analyze the convergence of this method and show that it is especially suitable for accelerating the convergence of vector sequences that are obtained when one solves linear systems of equations iteratively.

E-1328

*Technion - Israel Institute of Technology, Haifa, Israel and NRC-NASA Research Associate.

1. INTRODUCTION

Recently Smith and Ford [1] have reviewed and tested numerically four methods of convergence acceleration for sequences of vectors. These methods are the minimal polynomial extrapolation (MPE), the reduced rank extrapolation (RRE), and the scalar and vector epsilon algorithms. One of the conclusions of this report, based on numerical experimentation, is that the MPE and the RRE have about the same properties, and in general, have better convergance than the others.

The purpose of the present work is to develop a general framework within which one can define a large class of vector acceleration methods, the MPE and the RRE being in this class. We shall motivate this development in the way Shanks [3] motivated his development of the e_k-transformation for scalar sequences.

Shanks starts with a scalar sequence S_m, $m = 0, 1, \ldots$, that has the property

$$(1.1) \qquad S_m \sim S + \sum_{i=1}^{\infty} \alpha_i \lambda_i^m \quad \text{as } m \to \infty,$$

where S is $\lim_{m\to\infty} S_m$ if all $|\lambda_i| < 1$; otherwise S is called the "anti-limit" of the sequence S_m, $m = 0, 1, \ldots$ As one way of approximating S, Shanks proposes to solve the 2k+1 nonlinear equations

$$(1.2) \qquad S_m = S_{n,k} + \sum_{i=1}^{k} \tilde{\alpha}_i \tilde{\lambda}_i^m, \quad n \le m \le n+2k$$

for $S_{n,k}$, $\tilde{\alpha}_i, \tilde{\lambda}_i$, $i = 1, \ldots, k$. By solving the system in (1.2), Shanks has expressed $S_{n,k}$ as the quotient of two determinants as follows:

$$(1.3)\qquad S_{n,k} = \frac{\begin{vmatrix} S_n & S_{n+1} & \cdots & S_{n+k} \\ \Delta S_n & \Delta S_{n+1} & \cdots & \Delta S_{n+k} \\ \vdots & \vdots & & \vdots \\ \Delta S_{n+k-1} & \Delta S_{n+k} & \cdots & \Delta S_{n+2k-1} \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \Delta S_n & \Delta S_{n+1} & \cdots & \Delta S_{n+k} \\ \vdots & \vdots & & \vdots \\ \Delta S_{n+k-1} & \Delta S_{n+k} & \cdots & \Delta S_{n+2k-1} \end{vmatrix}}$$

Two equivalent formulations follow from (1.3):

a) It is clear from (1.3) that $S_{n,k}$ along with the k parameters β_i, i = 0, 1, . . . , k-1, solves the set of k+1 linear equations

$$(1.4)\qquad S_m = S_{n,k} + \sum_{i=0}^{k-1} \beta_i \, \Delta S_{m+i}, \quad n \le m \le n+k.$$

By taking the differences of the equations in (1.4), we can see that the β_i satisfy

$$(1.5)\qquad \Delta S_m = \sum_{i=0}^{k-1} \beta_i \, \Delta^2 S_{m+i}, \quad n \le m \le n+k-1,$$

and that once the β_i have been determined, $S_{n,k}$ can be computed from one of the equations in (1.4).

b) From (1.4) it follows that $S_{n,k}$ along with the parameters γ_i, i = 0, . . . , k, satisfy the equations

$$(1.6)\qquad S_{n,k} = \sum_{i=0}^{k} \gamma_i \, S_{m+i}, \quad n \le m \le n+k,$$

subject to

$$(1.7)\qquad \sum_{i=0}^{k} \gamma_i = 1.$$

By taking differences of the equations in (1.6), we can see that the γ_i satisfy

$$(1.8)\qquad 0 = \sum_{i=0}^{k} \gamma_i \, \Delta S_{m+i}, \quad n \le m \le n+k-1,$$

subject to (1.7). Once the γ_i have been determined, $S_{n,k}$ can be computed from one of the equations in (1.6). Furthermore, if $\gamma_k \neq 0$, then (1.7) and (1.8) are equivalent to

$$(1.9) \qquad \sum_{i=0}^{k-1} c_i \, \Delta S_{m+i} = -\Delta S_{m+k}, \; n \le m \le n+k-1,$$

where

$$(1.10) \qquad \gamma_i = \frac{c_i}{\sum_{j=0}^{k} c_j}, \; 0 \le i \le k-1, \; c_k \equiv 1,$$

ORIGINAL PAGE IS OF POOR QUALITY

provided $\sum_{j=0}^{k} c_j \neq 0.$

It has been proved by Wynn [4], that $S_{n,k}$, when applied to sequences S_m, m = 0, 1, . . . , that are of the form given in (1.1), converges to S as $n \to \infty$ (k fixed), under certain conditions on the λ_i, faster than S_n itself. Wynn actually gives rates of convergence, and also analyzes the stability properties of the approximations $S_{n,k}$, when errors are introduced in the S_m.

In the next section we shall extend the two formulations (a and b) given above for deriving Shanks' transformations, to vector sequences, and shall obtain a class of vector acceleration methods that includes some of the known methods as well as some new ones. In Section 3 we shall analyze one of these methods, subject to an assumption on the vector sequence analogous to (1.1), and shall prove that it accelerates the convergence of the sequence, and give an actual rate of convergence for it.

2. DERIVATION OF VECTOR ACCELERATION METHODS

Let us now consider a sequence of vectors, x_m, m = 0, 1, . . . , in a general space B, satisfying

$$x_m \sim s + \sum_{i=1}^{\infty} v_i \lambda_i^m \text{ as } m\to\infty, \tag{2.1}$$

where s and the v_i are vectors, and the λ_i are scalars. Again S is the limit or anti-limit of this sequence, depending on the λ_i. A very simple and practical example of such a sequence is that produced by a matrix iterative technique for solving the equation

$$x = Ax + b, \tag{2.2}$$

where A is an M x M matrix, and b and x are M-dimensional column vectors. If s is the solution to (2.2) and for given x_0, the vectors x_m are generated by

$$x_{m+1} = Ax_m + b, \quad m = 0, 1, \ldots, \tag{2.3}$$

then

$$x_m = s + \sum_{i=1}^{M} v_i \lambda_i^m, \quad m = 0, 1, \ldots, \tag{2.4}$$

where λ_i and v_i are the eigenvalues and corresponding eigenvectors respectively, of the matrix A, assuming that A has precisely M eigenvectors. The condition stated in (2.1) is analogous to that stated in (1.1) for scalar sequences. Since the Shanks' transformation accelerates the convergence of scalar sequences satisfying (1.1), we expect that its extensions to the vector case, through the formulations a) and b) in the previous section, will also produce acceleration of convergence for vector sequences satisfying (2.1).

The extensions of the two formulations can be achieved as follows:

a) Let us start by writing (1.3)-(1.4) in terms of the vector sequence. We have

$$S_{n,k} = x_n - \sum_{i=0}^{k-1} \beta_i \Delta x_{n+i} \tag{2.5}$$

with β_i obtained from the overdetermined system

(2.6) $$\Delta x_m = \sum_{i=0}^{k} \beta_i \, \Delta^2 x_{m+i}, \quad n \le m \le n+k-1$$

by some technique.

b) If we write (1.6)-(1.10) in terms of the vector sequence, we find

(2.7) $$S_{n,k} = \sum_{i=0}^{k} \gamma_i \, x_{n+i},$$

where γ_i are obtained from the overdetermined system

(2.8) $$\sum_{i=0}^{k-1} c_i \, \Delta x_{m+i} = -\Delta x_{m+k}, \quad n \le m \le n+k-1,$$

and

(2.9) $$\gamma_i = \frac{c_i}{\sum_{j=0}^{k} c_j}, \quad 0 \le i \le k-1, \; c_k \equiv 1,$$

provided $$\sum_{j=0}^{k} c_j \neq 0.$$

We see that for both approaches, we need to "solve" an overdetermined and, in general, inconsistent system of equations of the form

(2.10) $$\sum_{i=0}^{k-1} d_i \, w_{m+i} = \tilde{w}_m, \quad n \le m \le n+k-1,$$

where w_j and $\tilde{w}_j$ are members of the space B and d_i are unknown scalars. If r, the dimension of B, is greater than or equal to k, then even one of the equations in (2.10) gives rise to an overdetermined system. We can, however, propose various ways for obtaining a set of d_i, that "solves" (2.10) in some sense. In what follows, we give three such methods, with the understanding that other methods can also be proposed.

1) Assuming $n \geq k$, solve the overdetermined system [eq. (2.10) with m=n only]

ORIGINAL PAGE IS OF POOR QUALITY

$$\sum_{i=0}^{k-1} d_i \, w_{n+i} = \tilde{w}_n \tag{2.11}$$

by least squares. For finite dimensional spaces B, this method gives us RRE for approach a), and it gives MPE for approach b).

2) Assuming $r \geq k$, solve the set of k equations

$$\sum_{i=0}^{k-1} d_i \, Q_j \, (w_{n+i}) = Q_j \, (\tilde{w}_n), \; j = 0, 1, \ldots, k-1, \tag{2.12}$$

where Q_j are linearly independent operators in the dual space of B. For Hilbert spaces, we can take $Q_j(y) = (q_j, y)$, where q_j are vectors and (. , .) is an inner product.

3) Solve the set of equations

$$\sum_{i=0}^{k-1} d_i \, Q(w_{m+i}) = Q(\tilde{w}_m), \; n \leq m \leq n+k-1, \tag{2.13}$$

where Q is an operator in the dual space of B. This method is similar to that introduced by Brezinski [2].

3. A CONVERGENCE RESULT

We now give a convergence result for the method that is obtained from approach b) using the procedure in 2) in the previous section. We shall state the result, but leave out its proof. The proof, along with several other new results will be included in the final version of this paper. For simplicity we shall assume that B is finite dimensional.

Theorem: Let the sequence x_m, m = 0, 1, . . . be as in (2.3)-(2.4). Let $S_{r,k}$ be obtained from (2.7)-(2.9) with the c_i obtained by solving the linear system

$$\sum_{i=0}^{k-1} c_i \, Q_j \, (\Delta x_{n+i}) = -Q_j(\Delta x_{n+k}), \; j = 0, 1, \ldots, k-1$$

where Q_j are k linearly independent operators in the dual space of B. (For example, $Q_j(y)$ = (j+1)st component of y, j = 0, 1, . . . , k-1.) If $|\lambda_1| \geq |\lambda_2| \geq . . . \geq |\lambda_k| > |\lambda_{k+1}| > |\lambda_{k+2}| \geq . . .$, then, in general,

$$(3.1) \qquad S_{n,k} - S = O(\lambda_{k+1}^n) \text{ as } n\to\infty,$$

Recall that $x_n - S = O(\lambda_1^n)$ as $n\to\infty$. Therefore, whenever the conditions of the theorem above are satisfied, then $S_{n,k}$ converges to S faster than x_n when the latter converges, and if $|\lambda_{k+1}| < 1$, then $S_{n,k}$ converges to S even if x_n may diverge.

REFERENCES

[1] D. A. Smith, W. F. Ford, Extrapolation methods for vector sequences. Presented at the American Mathematical Society, Jan. 1980.

[2] C. Brezinski, A general extrapolation algorithm, Numer. Math., 35 (1980), pp. 175-187.

[3] D. Shanks, Non-linear transformations of divergent and slowly convergent sequences, J. Math. and Physics, 34 (1955), pp. 1-42.

[4] P. Wynn, On the convergence and stability of the epsilon algorithm, SIAM J. Numer. Anal., 3 (1966), pp. 91-122.