EXTRAPOLATING TO THE LIMIT OF A VECTOR SEQUENCE[1]

R. P. Eddy

Computation, Mathematics,
and Logistics Department
David W. Taylor Naval Ship R&D Center
Bethesda, Maryland

Several computational schemes are presented for implementing a vector extrapolation formula which estimates the limit of a sequence generated by the linear iteration $x(n+1) = Ax(n)+b$, where x and b are m x 1 vectors and A is an m x m matrix. Essential use is made of the fact that this sequence involves a geometric series in the matrix A. The extrapolation formula generalizes the well known formula for scalar sequences known as the Aitken Δ^2 or the Shanks e_1 process. In general, for a linear iteration the successive difference vectors u(n) satisfy $u(n+1) = Au(n)$. Only if the matrix A possesses a single dominant eigenvalue λ, and if iteration has reached a stage where u(n) is proportional to the corresponding eigenvector so that $u(n+1) \approx \lambda u(n)$, can the Aitken-Shanks extrapolation formula be applied successfully. More generally, if iteration has reached a stage where u(n) can be expressed, to within the accuracy carried, as a linear combination of r eigenvectors associated with r dominant eigenvalues, then only r successive difference vectors can be linearly independent to within that accuracy. The appropriate extrapolation formula requires the generalized inverse of an m x r matrix of second difference vectors. Determination of the rank r is the crux of the problem and can be troublesome. The most reliable methods involve the singular value decomposition of one of the difference matrices; two methods of so doing are among those presented. A close connection between these methods and Wynn's vector epsilon algorithm is conjectured.

[1]This research was carried out under the Mathematical Sciences Research Program, Task Area SR0140301 (Math Sciences), supported by the Naval Sea Systems Command.

ISBN 0-12-734250-8

I. INTRODUCTION

The computational methods presented here were developed for the purpose of accelerating convergence of the sequence of vectors generated in the iterative solution, by whatever scheme, of a system of m simultaneous linear equations in m variables

$$x_{n+1} = A\, x_n + b \tag{1.1}$$

where x and b are m x 1 vectors and A is an m x m matrix. Large systems of this type occur in the finite difference or finite element solution of elliptic partial differential equations. Important problems of this nature arise, for example, in studying fluid flow about naval vessels and also about aircraft and missiles.

The methods described here constitute a genuine generalization to multivariable systems of the well known Aitken (1926) Δ^2 - Shanks (1955) e_1 extrapolation formula

$$\tilde{x} = (x_{n-1}x_{n+1} - x_n^2)/(x_{n+1} - 2x_n + x_{n-1}) \tag{1.2}$$

which was derived for scalar sequences. They are based upon what seems to be a new approach: the approximation of (I-A) appearing in the exact extrapolated solution of (1.1), by a generalized inverse of rank r ($1 \leq r < m$) constructed from matrices whose columns are difference vectors.

Several computational approaches are indicated; however, computational experience with them is still very limited and the computer programs which implement them are still in the experimental stage.

Key roles in the theory are played by several concepts from linear algebra, some of which may not yet be well known. They are:

(1) Matrix geometric series: $I + A + A^2 + \ldots$

(2) Rank of a rectangular matrix = number of linearly independent columns or rows.

(3) The generalized inverse of a rectangular matrix (Penrose (1955, 1956), Greville (1959, 1960), Boullion and Odell (1971), Ben-Israel and Greville (1974), Stewart (1973)).

(4) The approximation of a given matrix by another of lower rank, especially by use of the next concept.

(5) The singular value decomposition of a rectangular matrix (Eckart and Young (1936), Golub and Kahan (1965), Golub and Reinsch (1970), Ben-Israel and Greville (1974), Stewart (1973)).

Due to limitations of space, most derivations, proofs, and discussions of the several computational strategies indicated have been omitted. A more detailed exposition will be given in a forthcoming internal report.

II. PRELIMINARIES

From the vector iteration equation (1.1) it is trivial to show that the successive m-component difference vectors

$$u_n = x_{n+1} - x_n \tag{2.1}$$

satisfy

$$u_n = A\, u_{n-1} = \cdots = A^n\, u_0 \tag{2.2}$$

The second difference vectors are given by

$$h_n = u_{n+1} - u_n = (A-I)\, u_n \tag{2.3}$$

The following m x n ($1 \leq n \leq m$) rectangular matrices will be used repeatedly; their columns are first or second difference vectors, as indicated.

$$U = (u_0, u_1, \cdots u_{n-1}) \tag{2.4}$$

$$U_1 = (u_1, u_2, \cdots u_n) \tag{2.5}$$

$$H = (h_0, h_1, \cdots h_{n-1}) \tag{2.6}$$

These matrices are related by

$$U_1 = A\,U \tag{2.7}$$

$$H = U_1 - U = (A-I)U \tag{2.8}$$

The latter of these equations is of fundamental importance in the theory to be presented.

III. FULL RANK EXTRAPOLATION

If the sequence of vectors generated by (1.1) converges, then the limit, $\hat{x}$, is given by

$$\begin{aligned}\hat{x} = \lim_{n\to\infty} x_n &= \lim(x_0 + (x_1 - x_0) + (x_2 - x_1) + \cdots) \\ &= \lim(x_0 + u_0 + u_1 + u_2 + \cdots) \\ &= x_0 + \lim(I + A + A^2 + \cdots)u_0 = x_0 + (I - A)^{-1}u_0\end{aligned} \tag{3.1}$$

This expression is formally valid, whether the vector sequence converges or not, provided that no eigenvalue of A is +1, so that the indicated inverse exists.

If the matrices U and H are both m x m and if H is non-singular, then (2.8) yields

$$(I - A)^{-1} = -U\,H^{-1} \tag{3.2}$$

and (3.1) becomes

$$\hat{x} = x_0 - U H^{-1} u_0 \tag{3.3}$$

or $$\hat{x} = x_0 + U\xi \tag{3.4}$$

$$0 = u_0 + H\xi \tag{3.5}$$

These equations constitute the full rank generalization to vector space of the Aitken (1926) Δ^2 - Shanks (1955) e_1 scalar extrapolation formula. Although theoretically (3.3) yields the exact limit, $\hat{x}$, it is not really useful for solving linear systems because it requires solving a linear system of the same order as the original. However, it does provide a pattern for the more useful reduced rank extrapolation to be taken up next.

IV. REDUCED RANK EXTRAPOLATION

Reduced rank, or rank r extrapolation, where $1 \leq r < m$, applies when only r successive vectors x_n (or u_n or h_n) can be linearly independent, at least to computational accuracy. This situation obtains when the iteration (1.1) has reached the point at which the x_n can be expressed as linear combinations of only r (instead of all m) eigenvectors (or principal vectors in case of degeneracy) belonging to the r eigenvalues of A having greatest absolute value. Contributions associated will all of the remaining m - r eigenvalues of lesser magnitude are then too small to be significant. These latter assertions are proved by combining (2.2) with the transformation which takes the iteration matrix A into classical diagonal (or Jordan) canonical form.

Determination of the rank r may be a delicate problem computationally. This problem will be taken up as part of the discussion of computational strategies in section 5.

Under the conditions considered here, not enough information is available to determine $(I - A)^{-1}$ exactly from (2.8) since U and H can have only r linearly independent columns. The best that can be done is to approximate $(I - A)^{-1}$ by a matrix of rank r, where $1 \leq r < m$, obtained from (2.8), which leads to a generalized inverse.

In (3.1), let the exact limit $\hat{x}$ be replaced by the extrapolated $\tilde{x}$ and let

$$(I - A)^{-1} u_0 = U\xi$$

From (4.1) and (2.8) it follows that

$$u_0 = (I - A)U\xi = - H\xi \tag{4.2}$$

so that the rank r extrapolation is given by

$$\tilde{x} = x_0 + U\xi \tag{4.3}$$

$$0 = u_0 + H\xi \tag{4.4}$$

Solve (4.4) by the method of least squares

$$0 = H^* u_0 + (H^*H)\xi \tag{4.5}$$

whence

$$\xi = - H^+ u_0 \tag{4.6}$$

where

$$H^+ = (H^*H)^{-1} H^* \tag{4.7}$$

is the generalized inverse of H. The rank r extrapolation formula becomes

$$\tilde{x} = x_0 - U H^+ u_0$$

These matrices have the following dimensions: U and H (m x r), H^+ (r x m), x, x_0, and u_0 (m x 1).

It is also possible to solve (4.4) for ξ by minimizing the maximum norm (Chebyshev approximation); the consequences of this approach have not yet been worked out.

Passing notice should be taken of the special case r=1. The matrices U and H then consist of a single column each and the extrapolation formula (4.8) collapses to the well known successive over-relaxation (SOR) formula

$$\tilde{x} = x_0 + \omega(x_1 - x_0) \tag{4.9}$$

with the over-relaxation factor ω prescribed:

$$\omega = - h_0^* u_0 / h_0^* h_0 \tag{4.10}$$

It might be expected that the same result would be given by the vector epsilon algorithm. (Wynn (1962), Brezinski (1977)). But application of the lozenge rule $(u_1^+ - u_0^+)(\varepsilon_2^0 - x_1) = 1$ yields

$$\varepsilon_2^0 = x_0 + (u_1^+ - u_0^+)^+ u_1^+ u_0 \neq \tilde{x} \tag{4.11}$$

The error in the extrapolated vector $\tilde{x}$ can be shown to be related to the error of x_0 by

$$\tilde{x} - \hat{x} = (I - U U^+)(x_0 - \hat{x}) \tag{4.12}$$

Geometrically, this says that $\tilde{x} - \hat{x}$ is orthogonal to the projection of $x_0 - \hat{x}$ onto the subspace spanned by the r columns of U.

A certain covariance property, generalizing that given by Shanks (1955) for his scalar transforms e_k^n, is easily demonstrated. Let all x_n undergo the linear transformation $x \to C x + f$, where C is an m x m matrix and f is a vector. Then

$$\tilde{x} \to C(x_0 - U(H^*C^*CH)^{-1} H^*C^*C u_0) + f \tag{4.13}$$

Thus the extrapolated vector undergoes the same linear transformation $\tilde{x} \rightarrow \tilde{x} + f$ provided that either $C^*C = c^2I$ or $C = cI$ holds, where c is a scalar and I is the identity.

V. COMPUTATIONAL STRATEGIES

Two interrelated problems must be attacked: (1) how to calculate H^+ in (4.8) or, equivalently, how to calculate ξ in (4.4), and (2) how to determine the rank r of H (or U). Since the rank may be very poorly defined, it seems best to concentrate on the most refined technique available using the singular value decomposition (SVD) of H or U as developed by Golub and collaborators. (Golub and Kahan (1965), Golub and Reinsch (1970), Stewart (1973), Fortran Algorithm 358 by Businger and Golub (1969)). Only passing mention is made of two other possible techniques: (1) the triangular square root (Choleski) method of solving the normal equations (4.5), increasing the dimension with each iteration until a diagonal element becomes too small, and (2) stepwise increase in dimension of H^+, as described by Greville (1960), until some h_n becomes linearly dependent upon its predecessors.

The most straightforward SVD method consists of cyclic n+1 step iteration with n determined by storage capability. At the end of each cycle the following steps are performed: construct H (n columns), obtain its singular value decomposition, and determine its rank, r. If r=n, use rank n extrapolation and recycle; if $r<n$, use rank r extrapolation and stop. The singular value decomposition is $H = W\ D\ V^*$, $H^+ = V\ D^+ W^*$ where W (m x n) and V (n x n) are unitary and D (n x n) is diagonal and non-negative. Rank $r \leq n$ is the

number of elements of D which exceed a threshold determined by the number of significant digits in a computer word. To obtain D^+, replace each of these r elements by its reciprocal, the rest by zero.

An elegant alternative is the following approach. Rewrite (4.4) as

$$\bar{U}(F\xi - f) = 0 \tag{5.1}$$

$$\bar{U} = (u_0, u_1, \cdots u_n) \tag{5.2}$$

F ((n+1) x n) is the column differencing operator, and f is a vector having first element unity, the remaining n being zero. Obtain the singular value decomposition $\bar{U} = W D V^*$ and determine the rank r. It is necessary that $r < n+1$. Let V_0 ((n+1) x (n+1-r)) consist of those columns of V which belong to zero singular values, let s be the n+1-r component vector of column sums of V_0, and let

$$w = V_0 \, s/(s^*s) \tag{5.3}$$

Then, from (4.3), the extrapolated vector is

$$\tilde{x} = \sum_{k=0}^{n} x_k \, w_{k+1} \tag{5.4}$$

The possibility to carry out explicity all steps except obtaining the SVD follows from the fact that F^+ can be exhibited explicitly and simply.

REFERENCES

Aitken, A.C. (1926). Proc. Roy. Soc. Edinburgh A46:289-305.
Ben-Israel, A., and Greville, T.N.E. (1974). "Generalized Inverses," John Wiley & Sons, New York.
Boullion, T.L., and Odell, P.L. (1971). "Generalized Inverse Matrices," Wiley-Interscience, New York.
Brezinski, C. (1977). "Accélération de la Convergence en Analyse Numérique," (Lecture Notes in Math #584), Springer-Verlag, Berlin, Heidelberg, New York.
Businger, P.A., and Golub, G.H. (1969). Comm. ACM 12:564-565.
Eckart, C., and Young, G. (1936). Psychometrika 1:211-218.
Golub, G.H., and Kahan, W. (1965). SIAM J. Num. Analysis 2: 205-224.
Golub, G.H., and Reinsch, C. (1970). Num. Math. 14:403-420; also in Wilkinson, J.H., and Reinsch, C. (Eds.) (1971), 134-151.
Greville, T.N.E. (1959). SIAM Rev. 1:38-43.
Greville, T.N.E. (1960). SIAM Rev. 2:15-22.
Penrose, R. (1955). Proc. Cambridge Phil. Soc. 51:406-413.
Penrose, R. (1956). Proc. Cambridge Phil. Soc. 52:17-19.
Shanks, D. (1955). J. Math. Phys. (MIT) 34:1-42.
Stewart, G.W. (1973). "Introduction to Matrix Computations," Academic Press, New York.
Wilkinson, J.H., and Reinsch, C. (Eds.) (1971). "Linear Algebra," Springer-Verlag, New York, Heidelberg, Berlin.
Wynn, P. (1962). Math. Comp. 16:301-322.