COMMUNICATIONS IN APPLIED NUMERICAL METHODS, Vol. 2. 385–392 (1986)

ACCELERATION OF VECTOR SEQUENCES BY MULTI-DIMENSIONAL Δ^2 METHODS

ALLAN J. MACLEOD

Department of Management Studies, Scottish College of Textiles, Galashiels, Scotland

SUMMARY

Nine different algorithms for the acceleration of vector sequences are considered. All have the property that, in one dimension, they reduce to the standard Aitken Δ^2 process. Different rules for deciding when to accelerate are also discussed. Theoretical and practical results show a wide range of performance, but it is shown that certain methods can be recommended for use.

INTRODUCTION

Vector sequences occur in a great many numerical problems. In many cases, however, the rate of convergence is poor, leading one to look for methods which will accelerate the convergence.

In one dimension the most well-known method of acceleration is Aitken's Δ^2 process. If $x_0, x_1, x_2, \ldots$ is a scalar sequence, then we accelerate by using the formula

$$x'_{i+2} = x_{i+2} - \frac{(x_{i+2} - x_{i+1})^2}{x_{i+2} - 2x_{i+1} + x_i} \tag{1}$$

$$= x_{i+2} - \frac{(\Delta x_{i+2})^2}{\Delta^2 x_{i+2}} \tag{2}$$

It has been shown by Pennacchi,[1] and Germain-Bonne[2] that, for linear convergence, the Aitken method is optimal in a variety of ways.

In this paper, various methods are considered for accelerating a sequence of vectors $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$, each of which has the property that, in one dimension, it reduces to the Aitken method.

METHODS

Method 1

If we are dealing with n-dimensional vectors, then the sequence of vectors can be considered as n independent scalar sequences, and the Aitken method can be applied to these. Thus, a new vector $\mathbf{x}'_{i+2}$ is generated according to the formula

$$(x'_{i+2})_j = (\mathbf{x}_{i+2})_j - \frac{\{(\mathbf{x}_{i+2})_j - (\mathbf{x}_{i+1})_j\}^2}{(\mathbf{x}_{i+2})_j - 2(\mathbf{x}_{i+1})_j + (\mathbf{x}_i)_j} \tag{3}$$

$$j = 1, 2, \ldots, n.$$

This method is used in Willmott and Kulakowski,[3] who discuss thermal regenerator simulations.

0748–8025/86/040385–08$05.00 *Received September 1985*

Method 2

In a paper on atomic field calculations, Roothan and Bagus[4] describe an acceleration method which they credit to R. A. Sack. The assumption used is that for successive iterates, the end-points of the vectors $\mathbf{x}_i$ describe a spiral in a plane centred on the limit $\mathbf{x}$, i.e. we assume

$$\mathbf{x}_{i+1} - \mathbf{x} = m\, R(\psi)\, (\mathbf{x}_i - \mathbf{x}) \,, \tag{4}$$

where m is a scale factor, and $R(\psi)$ is the matrix representing a two-dimensional rotation over ψ.

An expression for $\mathbf{x}$ can be derived which gives the acceleration formula

$$\mathbf{x}'_{i+2} = (\alpha - 2\beta + \gamma)^{-1}\, \{\alpha\, \mathbf{x}_i - 2\beta \mathbf{x}_{i+1} + \gamma \mathbf{x}_{i+2}\} \tag{5}$$

where $\alpha = \| \mathbf{x}_{i+2} - \mathbf{x}_{i+1}\|^2$, $\beta = (\mathbf{x}_{i+2} - \mathbf{x}_{i+1}).(\mathbf{x}_{i+1} - \mathbf{x}_i)$, and $\gamma = \|\mathbf{x}_{i+1} - \mathbf{x}_i\|^2$.

Expression (5) can be derived from a totally different set of assumptions, as follows. In 1962, Wynn[5] extended his scalar ϵ-algorithm to vector and matrix sequences. The method for vectors is defined by the equations

$$\begin{aligned} \boldsymbol{\epsilon}^{(i)}_{-1} &= \mathbf{0},\ \boldsymbol{\epsilon}^{(i)}_{0} = \mathbf{x}_i,\ i = 0, 1, 2, \ldots \\ \boldsymbol{\epsilon}^{(i)}_{k+1} &= \boldsymbol{\epsilon}^{(i)}_{k-1} + (\Delta \boldsymbol{\epsilon}^{(i)}_{k})^{-1},\ i, k = 0, 1, \ldots \end{aligned} \tag{6}$$

with the inverse $\mathbf{y}^{-1}$ of a vector $\mathbf{y}$ defined as

$$\mathbf{y}^{-1} = \mathbf{y}/\|\mathbf{y}\|^2$$

For $k = 2$, these formulae lead to the acceleration formula

$$\mathbf{x}'_{i+2} = \mathbf{x}_{i+1} + \frac{\gamma(\mathbf{x}_{i+2} - \mathbf{x}_{i+1}) - \alpha(\mathbf{x}_{i+1} - \mathbf{x}_i)}{\|\mathbf{x}_{i+2} - 2\mathbf{x}_{i+1} + \mathbf{x}_i\|^2} \tag{7}$$

and a little algebraic manipulation shows (7) to be identical with (5).

Method 3

In a paper on the numerical solution of non-linear integral equations, Anderson[6] derived a method, based essentially on the following argument:

Given $\mathbf{x}_i$, $\mathbf{x}_{i+1}$, $\mathbf{x}_{i+2}$, and a scalar μ, let

$$\begin{aligned} \mathbf{u} &= \mathbf{x}_{i+1} + \mu\, (\mathbf{x}_i - \mathbf{x}_{i+1}) \\ \mathbf{v} &= \mathbf{x}_{i+2} + \mu\, (\mathbf{x}_{i+1} - \mathbf{x}_{i+2}) \end{aligned} \tag{8}$$

The value of μ was chosen to minimise $\|\mathbf{v} - \mathbf{u}\|^2$. It is straightforward to show that

$$\mu = \frac{\mathbf{r}_{i+2} \cdot (\mathbf{r}_{i+2} - \mathbf{r}_{i+1})}{\|\mathbf{r}_{i+2} - \mathbf{r}_{i+1}\|^2} \tag{9}$$

with $\mathbf{r}_j = \mathbf{x}_j - \mathbf{x}_{j-1}$, $j = 1, 2, \ldots$. The vector $\mathbf{v}$ was then defined to be the accelerated vector $\mathbf{x}'_{i+2}$. This method is also discussed by Jennings,[7] who calls it the SDM method, and by Irons and Tuck.[8]

Method 4

In a recent paper, Zienkiewicz and Lohner[9] propose an alternative method which uses the acceleration formula

$$\mathbf{x}'_{i+2} = \mathbf{x}_{i+2} - \frac{\|\mathbf{x}_{i+2} - \mathbf{x}_{i+1}\|^2 (\mathbf{x}_{i+2} - \mathbf{x}_{i+1})}{(\mathbf{x}_{i+2} - \mathbf{x}_{i+1})\cdot(\mathbf{x}_{i+2} - 2\mathbf{x}_{i+1} + \mathbf{x}_i)} \tag{10}$$

Method 5

In the paper by Jennings,[7] a method similar to method 4 is also given, which Jennings call the FDM method. We have

$$\mathbf{x}'_{i+2} = \mathbf{x}_{i+2} - \frac{(\mathbf{x}_{i+2} - \mathbf{x}_{i+1}).(\mathbf{x}_{i+1} - \mathbf{x}_i)\,(\mathbf{x}_{i+2} - \mathbf{x}_{i+1})}{(\mathbf{x}_{i+1} - \mathbf{x}_i).(\mathbf{x}_{i+2} - 2\mathbf{x}_{i+1} + \mathbf{x}_i)} \tag{11}$$

Method 6

This method is, as far as can be ascertained, new and was communicated by Dr. D. W. Arthur of Edinburgh University.

Suppose that $\mathbf{e}_i = \mathbf{x}_i - \mathbf{x}$, $i = 0, 1, 2, \ldots$, with $\mathbf{x}$ the limit of the sequence, and suppose also that $\mathbf{e}_{i+1} = M\mathbf{e}_i$, where M is a symmetric matrix. Then it is easy to show that $\mathbf{e}_{i+1}.\mathbf{e}_{i+1} = \mathbf{e}_{i+2}.\mathbf{e}_i$, which can be rearranged to give

$$(\mathbf{x}_{i+2} - 2\mathbf{x}_{i+1} + \mathbf{x}_i).(\mathbf{x} - \mathbf{x}_{i+2}) = -\|\mathbf{x}_{i+2} - \mathbf{x}_{i+1}\|^2 \tag{12}$$

Equation (12) has an infinite number of solutions for $\mathbf{x} - \mathbf{x}_{i+2}$, but it can be shown that the solution with minimum norm is given by

$$\mathbf{x} - \mathbf{x}_{i+2} = \frac{-\|\mathbf{x}_{i+2} - \mathbf{x}_{i+1}\|^2}{\|\mathbf{x}_{i+2} - 2\mathbf{x}_{i+1} + \mathbf{x}_i\|^2} (\mathbf{x}_{i+2} - 2\mathbf{x}_{i+1} + \mathbf{x}_i) \tag{13}$$

If $\mathbf{x}$ is replaced by $\mathbf{x}'_{i+2}$ we have the required acceleration formula.

Method 7

This method is derived by assuming that eventually $\mathbf{e}_{i+1} = \lambda\mathbf{e}_i$. Thus

$$\begin{aligned} \mathbf{x} &= \mathbf{x}_i + (\mathbf{x}_{i+1} - \mathbf{x}_i) + (\mathbf{x}_{i+2} - \mathbf{x}_{i+1}) + \ldots \\ &= \mathbf{x}_i + (\mathbf{x}_{i+1} - \mathbf{x}_i)\,(1 + \lambda + \lambda^2 + \ldots) \\ &= \mathbf{x}_i + \frac{1}{1 - \lambda}(\mathbf{x}_{i+1} - \mathbf{x}_i) \end{aligned} \tag{14}$$

The value of λ is estimated by $|\lambda| = \dfrac{\|\mathbf{x}_{i+2} - \mathbf{x}_{i+1}\|}{\|\mathbf{x}_{i+1} - \mathbf{x}_i\|}$

with the sign of λ coming from the ratios of components of $(\mathbf{x}_{i+2} - \mathbf{x}_{i+1})$ and $(\mathbf{x}_{i+1} - \mathbf{x}_i)$.
If $\mathbf{x}$ is replaced by $\mathbf{x}'_{i+2}$, then the acceleration formula is

$$\mathbf{x}'_{i+2} = \mathbf{x}_i + \frac{1}{1-\lambda}(\mathbf{x}_{i+1} - \mathbf{x}_i) \tag{15}$$

This method is used in the paper by Orbach and Crowe.[10]

Method 8

If the telescoped series in Method 7 is started at $\mathbf{x}_{i+1}$ instead of $\mathbf{x}_i$, the following acceleration formula results:

$$\mathbf{x}'_{i+2} = \mathbf{x}_{i+2} - \frac{\lambda}{\lambda-1}(\mathbf{x}_{i+2} - \mathbf{x}_{i+1}) \tag{16}$$

Method 9

If we assume that $\mathbf{e}_{i+1} = \lambda\, \mathbf{e}_i$ eventually, then $\mathbf{e}_{i+2} = \lambda^2\mathbf{e}_i$. Thus

$$\mathbf{x} = \mathbf{x}_{i+2} - \lambda^2\mathbf{e}_i$$

But

$$\mathbf{x}_{i+1} - \mathbf{x}_i = \mathbf{e}_{i+1} - \mathbf{e}_i = (\lambda - 1)\mathbf{e}_i$$

Hence

$$\mathbf{x} = \mathbf{x}_{i+2} - \lambda^2\,(\mathbf{x}_{i+1} - \mathbf{x}_i)/(\lambda-1)$$

This gives the acceleration formula

$$\mathbf{x}'_{i+2} = \mathbf{x}_{i+2} - \frac{\lambda^2}{\lambda-1}(\mathbf{x}_{i+1} - \mathbf{x}_i) \tag{17}$$

ANALYSIS OF METHODS

Clearly, a full analysis of the proposed methods would be difficult and lengthy. It is possible, however, to gain some insight by assuming that $\mathbf{e}_{i+1} = M\mathbf{e}_i$, where M is a matrix with eigenvalues λ_1, $\lambda_2, \ldots, \lambda_n$ (in decreasing order of magnitude), and corresponding eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$. Assume that the eigenvectors form a basis.

Method 1 assumes that the matrix M is diagonal, ignoring the coupling which exists between equations. This can lead to severe problems and even divergence. This has been well known for many years (see Henrici[11]). The other methods all consider the vector as a single entity. They can be analysed using the technique used by Jennings.

Suppose we have three iterates $\mathbf{x}_i$, $\mathbf{x}_{i+1}$, $\mathbf{x}_{i+2}$, with $\mathbf{e}_{i+1} = M\mathbf{e}_i$, and $\mathbf{e}_{i+2} = M\mathbf{e}_{i+1}$. Let

$$\mathbf{e}_i = \sum_{j=1}^{n} a_j\, \mathbf{u}_j$$

Then

$$\mathbf{e}_{i+1} = \sum_{j=1}^{n} \lambda_j\, a_j\, \mathbf{u}_j, \text{ and } \mathbf{e}_{i+2} = \sum_{j=1}^{n} \lambda_j^2\, a_j\, \mathbf{u}_j$$

Let

$$\mathbf{e}'_{i+2} = \mathbf{x}'_{i+2} - \mathbf{x} = \sum_{j=1}^{n} \mu_j\, a_j\, \mathbf{u}_j$$

It is straightforward to derive formulae for the μ_j for Methods 2–9. These are given in Table I, where $A = \Sigma\, z_j$, $B = \Sigma\lambda_j z_j$, $C = \Sigma\lambda_j^2\, z_j$, with $z_j = (1 - \lambda_j)^2\, a_j^2$.

Table I. Scale factors for Methods 2–9

Method	μ_j
2	$(A\lambda_j^2 - 2B\lambda_j + C)/(A - 2B + C)$
3	$((A-B)\lambda_j^2 + (C-B)\lambda_j)/(A - 2B + C)$
4	$(B\lambda_j^2 - C\lambda_j)/(B - C)$
5	$(A\lambda_j^2 - B\lambda_j)/(A - B)$
6	$((A-2B)\lambda_j^2 + 2C\lambda_j - C)/(A - 2B + C)$
7	$(\sqrt{(A)}\lambda_j - \sqrt{(C)})/(\sqrt{(A)} - \sqrt{(C)})$
8	$(\sqrt{(A)}\lambda_j^2 - \sqrt{(C)}\lambda_j)/(\sqrt{(A)} - \sqrt{(C)})$
9	$((A-\sqrt{(AC)})\ \lambda_j^2 + (\lambda_j-1)C)/(A - \sqrt{(AC)})$

The components μ_j are essentially scale factors which must be compared with the unaccelerated components λ_j^2. If we assume M to be symmetric and positive definite, then it is possible to make some comparisons. In such a case, the eigenvalues of M will all be in the range 0 to 1.

The most influential component on the rate of convergence is that corresponding to $\mathbf{u}_1$, since λ_1 is the largest eigenvalue. It is easy to show that for Methods 2–9, $\mu_1 > 0$, and, for all but Method 7, that $\mu_1 < \lambda_1^2$, as we would like.

Method 7 is clearly unsuitable from another point of view. The unaccelerated scale function is λ_j^2, and all the method scale functions are also quadratics in λ_j, except Method 7.

To compare the other methods, consider the shape of the scale functions as λ_j varies from 0 to 1. Methods 3, 4, 5 and 8 have functions which are 0 at $\lambda_j = 0$, and 0 for λ_j in (0, 1). Methods 6 and 9 have a single zero for λ_j in (0, 1). Method 2 has a function which is always positive with a turning point in (0, 1).

This last statement suggests that Method 2 might not be so effective, since the other methods can reduce components to zero, whereas Method 2 cannot. Method 6 has a scale function which can be either convex (as all the rest are) or concave. If it was concave, the components corresponding to small eigenvalues would be expanded not contracted.

Thus the analysis implies that Methods 3, 4, 5, 8, and possibly 9, should be the most successful.

For M not positive definite it is impossible to state how the components will be affected in general.

POINT OF ACCELERATION

Along with the acceleration formula we require a rule to decide when to apply the formula. The most common method is to start with an initial approximation, iterate twice, accelerate, and repeat the process. This procedure can fail to converge, however, as was recognized by Willmott and Kulakowski.[3] They accelerated after every six iterations. This, however, was a good value for their particular problem. What we would like is a somewhat more automatic rule applicable to any problem.

The following method provides such a rule. It is derived in terms of a scalar sequence and can thus be applied to the norms of vectors.

If $x_{i+1} = g(x_i)$, then ultimately $x_{i+1} - x = \alpha(x_i - x)$, where α is the value of g' at the limit x. We can estimate α from $(x_{i+2} - x_{i+1})/(x_{i+1} - x_i)$. Alternatively, we could estimate it from $(x_{i+3} - x_{i+2})/(x_{i+2} - x_{i+1})$. Thus close to the limit $(x_{i+3} - x_{i+2})(x_{i+1} - x_i)/(x_{i+2} - x_{i+1})^2$ should be approximately 1. Call this ratio θ.

Then a possible decision rule is: wait till $|\theta - 1| <$ specified value, then iterate. The smaller the specified value the more iterations which must be performed before applying the acceleration formula.

NUMERICAL RESULTS

To assess the effectiveness of the methods, a simulation experiment was performed. Two different numerical problems were considered: (a) the solution of a set of linear equations using the

Gauss–Seidel method—the matrix and right-hand sides were randomly generated with the matrix being symmetric and diagonally-dominant; (b) the largest eigenvalue of a random symmetric matrix was found using the power method.

Both problems were repeated 1000 times, with the iteration being stopped when $\|\mathbf{x}_{i+1} - \mathbf{x}_i\| < \text{TOL}\,\|\mathbf{x}_{i+1}\|$. Seven decision rules were considered and for each combination of method and rule the average number of iterations to convergence calculated. Two values of TOL were used and the results are contained in Tables II–V. Also included is the average number of iterations needed when no acceleration was used.

The seven decision rules were:

(a) accelerate every two iterations,
(b) accelerate only if $|\theta-1| < 0{\cdot}1$,
(c) perform first acceleration when $|\theta - 1| < 0{\cdot}1$, then accelerate every two iterations,
(d) as (b), but with $|\theta - 1| < 0{\cdot}5$,
(e) as (c), but with $|\theta - 1| < 0{\cdot}5$,
(f) as (b), but with $|\theta - 1| < 1{\cdot}0$,
(g) as (c), but with $|\theta - 1| < 1{\cdot}0$.

For the Gauss–Seidel method we have the relation $\mathbf{e}_{i+1} = M\mathbf{e}_i.M$, however, will not be symmetric so the analysis performed on the methods might not hold. The relation is not true for the power method.

The results in Tables II–V clearly show Methods 1 and 7 to be very poor and they should never be used. Method 2 does not perform well on the power method, although it does better on Gauss–Seidel. Method 6 and Method 9 do fairly well, with Method 9 good for the power method.

Table II. Gauss–Siedel
Average number of iterations for TOL = 0·001
With no acceleration average = 6·17

Method	Decision rule (a)	(b)	(c)	(d)	(e)	(f)	(g)
1	9·35	6·07	6·10	6·11	6·27	6·27	6·59
2	5·40	6·01	6·00	5·78	5·73	5·65	5·53
3	4·98	6·00	5·98	5·68	5·62	5·42	5·30
4	6·46	6·03	6·04	5·90	5·93	5·95	6·04
5	5·08	6·01	5·99	5·71	5·65	5·49	5·37
6	5·99	6·03	6·03	5·99	5·81	5·89	5·70
7	11·21	6·09	6·17	6·25	6·65	6·60	7·50
8	5·46	6·01	5·99	5·79	5·72	5·68	5·53
9	6·00	6·03	6·02	5·88	5·82	5·87	5·73

Table III. Gauss–Siedel
Average number of iterations for TOL = 0·00001
With no acceleration average = 9·27

Method	Decision rule (a)	(b)	(c)	(d)	(e)	(f)	(g)
1	13·93	8·77	8·94	8·97	9·69	9·44	10·62
2	8·63	8·68	8·59	8·48	8·24	8·40	8·11
3	7·78	8·69	8·55	8·43	8·13	8·16	7·87
4	9·71	8·71	8·72	8·65	8·82	8·90	9·23
5	7·87	8·69	8·57	8·45	8·22	8·23	8·00
6	9·36	8·81	8·76	8·86	8·81	8·97	8·97
7	18·67	8·98	9·47	9·51	11·37	10·54	13·66
8	8·27	8·69	8·61	8·51	8·39	8·44	8·33
9	9·60	8·79	8·75	8·82	8·84	8·91	9·09

Table IV. Power method
Average number of iterations for TOL = 0·001
With no acceleration average = 7·94

Method	Decision rule (a)	(b)	(c)	(d)	(e)	(f)	(g)
1	12·61	7·29	8·37	7·39	9·17	8·44	10·54
2	6·67	6·96	6·83	6·87	6·64	6·95	6·63
3	5·34	6·55	6·42	6·33	5·99	6·38	5·65
4	18·80	7·05	8·40	7·02	9·99	7·46	14·65
5	5·41	6·53	6·41	6·32	5·98	6·39	5·63
6	6·01	6·61	6·54	6·33	6·13	6·23	5·95
7	10·28	7·47	7·28	7·51	7·45	7·78	8·54
8	6·06	6·73	6·61	6·53	6·17	6·48	5·81
9	5·79	6·51	6·45	6·21	6·01	6·05	5·65

Table V. Power method
Average number of iterations for TOL = 0·00001
With no acceleration average = 12·23

Method	Decision rule (a)	(b)	(c)	(d)	(e)	(f)	(g)
1	22·24	11·54	16·95	12·01	18·37	13·70	20·06
2	11·40	10·48	10·27	10·59	10·55	10·89	11·07
3	8·51	9·80	8·98	9·72	8·79	9·52	8·67
4	40·38	10·46	13·59	10·82	16·12	11·26	22·98
5	8·66	9·82	9·02	9·75	8·81	9·69	8·79
6	9·29	9·67	9·00	9·58	8·86	9·45	8·98
7	16·06	10·67	10·88	10·85	12·03	11·22	14·12
8	8·59	9·70	8·68	9·54	8·37	9·43	8·29
9	9·00	9·40	8·57	9·31	8·44	9·07	8·64

The major surprise is the very poor results for Method 4. It is not at all clear why this should be so, but inherent ill-conditioning might be a prime suspect.

The other methods, 3, 5 and 8, do well throughout, with Method 3 being, perhaps, the most consistently good, although no method has a clear supremacy.

As for decision rules, rules (a) and (g) seem to be best when used with a good acceleration formula. This is clearly a problem which needs more work done on it.

ACKNOWLEDGEMENTS

The author would like to thank Dr. D. W. Arthur of the University of Edinburgh for many helpful comments and advice which has helped the author greatly. The comments of a Referee are also gratefully acknowledged.

REFERENCES

1. R. Pennacchi, 'Le transformazioni razionali di una successione', *Calcolo*, **5**, 37–50 (1968).
2. B. Germain-Bonne, 'Estimation de la limite des suites et formalisation des procedes d'acceleration de convergence', *Thesis*, Univ. of Lille (1978).
3. A. J. Willmott and B. Kulakowski, 'Numerical acceleration of thermal regenerator simulations', *Int. j. numer. methods eng.*, **11**, 533–551 (1977).
4. C. C. J. Roothan and P. S. Bagus, 'Atomic self-consistent field calculations', in *Methods in Computational Physics*, vol. 2, Academic Press, New York, 1963.
5. P. Wynn, 'Acceleration techniques for iterated vector and matrix problems', *Maths. Comp.*, **16**, 301–322 (1963).
6. D. G. Anderson, 'Iterative procedures for non-linear integral equations', *J.A.C.M.*, **12**, 547–560 (1965).
7. A. Jennings, 'Accelerating the convergence of matrix iterative processes', *J. Inst. Maths Applics.*, **8**, 99–110 (1971).
8. B. M. Irons and R. C. Tuck, 'A version of the Aitken accelerator for computer iteration', *Int. j. numer methods eng.*, **1**, 275–277 (1969).

9. O. C. Zienkiewicz and R. Lohner, 'Accelerated relaxation or direct solution. Future prospects for FEM', *Int. j. numer. methods eng.*, **21**, 1–11 (1985).
10. O. Orbach and C. M. Crowe, 'Convergence promotion in the simulation of chemical processes with recycle—the dominant eigenvalue method', *Can. J. Chem. Eng.*, **49**, 509–513 (1971).
11. P. Henrici, *Elements of Numerical Analysis*, Wiley, New York, 1964.