REPORT

Haifa 1985 Conference on Matrix Theory

Abraham Berman
Department of Mathematics
Technion – Israel Institute of Technology
Haifa 32000, Israel

Yair Censor
Department of Mathematics and Computer Science
University of Haifa
Mount Carmel, Haifa 31999, Israel
and
Hans Schneider
Department of Mathematics
The University of Wisconsin
Madison, Wisconsin 53706
and
Technion – Israel Institute of Technology
Haifa 32000, Israel

Submitted by Hans Schneider

INTRODUCTION

The Haifa 1985 Conference is the second in a series (hopefully of length greater than 2) of conferences on matrix theory. It was held from 29–31 December, 1985 under the auspices of the Israel Mathematical Union, and hosted by the Department of Mathematics at the Technion and the Department of Mathematics and Computer Science at the University of Haifa. Participants from Israeli universities, research institutes, and high-tech industries gathered to listen to twenty-seven speakers, including three guests from the U.S.A. and West Germany. The informal exchange of information and ideas culminated in an open problem session.

The social program included receptions at the University of Haifa and at the Technion, and a guided tour of the Reuben and Edith Hecht Museum at the University of Haifa.

This report contains synopses of talks presented at the meeting, which were made available to us. They are arranged in alphabetical order. In coauthored synopses, the speakers' names are starred.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/86/$3.50

The program was as follows:

29 December
Technion, The Silver Institute of Biomedical Engineering Auditorium

10:00–11:00

Dimension of faces generated by certain positive linear operators, by R. Loewy, Technion.

Sign patterns of matrices and their inverses, by M. Berger, Weizmann Institute.

Chairman: A. Berman, Technion.

11:30–12:30

Submultiplicativity and mixed submultiplicativity of matrix norms and operator norms, by M. Goldberg, Technion.

The resolvent condition and uniform power boundedness, by E. Tadmor, Tel-Aviv University.

Chairman: H. Wolkowicz, University of Delaware.

14:00–15:00

An efficient preconditioning algorithm and its analysis, by I. Efrat, IBM Scientific Center, Haifa.

Parallel algorithms for triangular systems, by A. Lin, Technion.

Chairman: I. Gohberg, Tel-Aviv University.

15:30–16:30

Minimality and irreducibility of time-invariant boundary-value systems, by L. Lerer, Technion.

Maximum-entropy extensions of matrices and related problems, by I. Gohberg, Tel-Aviv University.

Chairman: I. Cederbaum, Technion.

16:30–17:00

Open-problems session

Chairman: D. Hershkowitz, Technion.

30 December
University of Haifa, Main Building, Room 608

9:00–10:30

Class functions of finite groups, nonnegative matrices, and generalized circulants, by D. Chillag, Technion.

On linearly constrained entropy maximization, by Y. Censor, University of Haifa.

Some theorems in matrix theory using optimization, by H. Wolkowicz, University of Delaware.

Chairman: J. Sonn, Technion.

11:00–12:30

Holdability, irreducibility, and M-matrices, by A. Berman, Technion.

On the uniqueness of the Lyapunov scaling factors, by D. Shasha, Technion.

On positive reciprocal matrices, by V. Mehrmann, University of Bielefeld.

Chairman: H. Schneider, University of Wisconsin and Technion.

14:00–15:30

On biholomorphic automorphisms of the unit ball of unitary matrix spaces, by J. Arazy, University of Haifa.

Perturbed and mixed Toeplitz matrices as generalization of the resultant matrix, by B. Kon, Technion.

Bezoutian for several matrix polynomials and polynomial Lyapunov-type equations, by M. Tismenetsky, IBM Scientific Center, Haifa.

Chairman: G. Moran, University of Haifa.

16:00–17:00

Powers of a nonnegative definite matrix related to interpolation by radial functions, by N. Dyn, Tel-Aviv University.

Using Gauss-Jordan elimination to compute the index, null space, and Drazin inverse, by U. Rothblum, Technion.

Chairman: A. Pinkus, Technion.

31 December

Technion, The Silver Institute of Biomedical Engineering Auditorium

9:30–10:30

Higher dimensional Euclidean and hyperbolic matrix spaces, by B. Schwarz, Technion.

From the complex numbers to complex matrices along the projective line, by A. Zaks, Technion.

Chairman: D. London, Technion.

11:00–12:00

Maximum-distance problem and band sequences, by A. Ben-Artzi, Tel-Aviv University.

Block-Hankel-matrix inversion and the partial-realization problem, by P. A. Fuhrmann, Ben-Gurion University at the Negev.

Chairman: V. Mehrmann, University of Bielefeld.

13:30–14:30

Eigenstructures and signal processing, by A. Bruckstein, Technion.
The structure of root clustering criteria, by S. Gutman, Technion.

Chairman: P. A. Fuhrmann, Ben-Gurion University at the Negev.

15:00–16:00

Matrices with sign symmetric diagonal shifts, by D. Hershkowitz, Technion.

Equality classes of matrices: The extremal case of an inequality due to Ostrowski, by H. Schneider, University of Wisconsin and Technion.

Chairman: Y. Censor, University of Haifa.

Synopses of the talks are presented below.

COMPUTING THE INDEX AND DRAZIN INVERSE USING THE SHUFFLE ALGORITHM

by KURT M. ANSTREICHER[1] and URIEL G. ROTHBLUM[2]*

The well-known Gauss-Jordan elimination procedure computes the inverse of a nonsingular matrix A by executing elementary row operations on the pair (A, I) to transform it into (I, A^{-1}). Moreover, Gauss-Jordan elimination can be used to determine whether or not a matrix A is nonsingular, in the case where this fact is not known *a priori*. We adapt the Gauss-Jordan elimination procedure via "shuffles" to obtain an algorithm which computes the index of a given matrix A and determines bases of the null spaces of the powers of A. In the worst case the algorithm requires less than $2n^3$ arithmetic operations, compared to the well-known bound of n^3 operations for the work needed to invert a nonsingular matrix using Gauss-Jordan elimination. Moreover, the procedure suggests adaptation of efficient techniques for computing inverses

[1]School of Organization and Management, Yale University, Box 1A, New Haven, CT 06520.

[2]Faculty of Industrial Engineering and Management, Technion–Israel Institute of Technology, Technion City, Haifa 32000, Israel. Research of this author was supported by Grant No. ECS-83-10213 from the National Science Foundation.

(cf. [13]). Finally, the Drazin inverse of the underlying matrix A has a simple representation in terms of the output of the algorithm and the matrix A itself.

An algorithm incorporating shuffle operations was first devised by Luenberger [9], who applied it to the study of singular difference equations. Anstreicher [1] analyzed Luenberger's algorithm as applied to the solution of singular systems of linear differential equations with constant coefficients. Previous methods for solving the latter used Drazin inverses (e.g., [3] and [12]), suggesting a connection between the shuffle algorithm and the Drazin inverse. These methods are closely related to the study of matrix pencils (see [11] and references therein).

To introduce our new algorithm we consider an example (cf. [2, p. 132]). Let

$$A = \begin{pmatrix} 2 & 4 & 6 & 5 \\ 1 & 4 & 5 & 4 \\ 0 & -1 & -1 & 0 \\ -1 & -2 & -3 & 3 \end{pmatrix}.$$

Elementary row operations transform (A, I) into

$$\begin{pmatrix} 1 & 0 & 1 & 0 & \frac{4}{3} & -\frac{5}{3} & -\frac{4}{3} & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & -\frac{1}{3} & \frac{2}{3} & \frac{4}{3} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{2}{3} & 1 \end{pmatrix}.$$

It is clear that A is not invertible, as a linear combination of its rows vanishes. The "shuffle step" will next exchange row(s) of zeros with the corresponding row(s) of the right-hand matrix. This yields

$$\begin{pmatrix} 1 & 0 & 1 & 0 & \frac{4}{3} & -\frac{5}{3} & -\frac{4}{3} & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & -\frac{1}{3} & \frac{2}{3} & \frac{4}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{2}{3} & 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

One then resumes elementary row operations, which result in

$$\begin{pmatrix} 1 & 0 & 1 & 0 & \frac{4}{3} & -\frac{5}{3} & -\frac{4}{3} & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & -\frac{1}{3} & \frac{2}{3} & \frac{4}{3} & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{3} & -\frac{1}{3} & -\frac{5}{3} & 0 \end{pmatrix}.$$

A second shuffle is next performed, yielding

$$\begin{pmatrix} 1 & 0 & 1 & 0 & \frac{4}{3} & -\frac{5}{3} & -\frac{4}{3} & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & -\frac{1}{3} & \frac{2}{3} & \frac{4}{3} & 0 \\ -\frac{1}{3} & -\frac{1}{3} & -\frac{5}{3} & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Elementary row operations are now finally used to convert the left-hand matrix to the identity, yielding

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \frac{16}{9} & -\frac{20}{9} & -\frac{19}{9} & 0 \\ 0 & 1 & 0 & 0 & \frac{4}{9} & -\frac{5}{9} & -\frac{16}{9} & 0 \\ 0 & 0 & 1 & 0 & -\frac{4}{9} & \frac{5}{9} & \frac{7}{9} & 0 \\ 0 & 0 & 0 & 1 & -\frac{1}{3} & \frac{2}{3} & \frac{4}{3} & 0 \end{pmatrix}.$$

at which point the algorithm terminates.

Our main results are as follows. First, the algorithm always terminates in a finite number of shuffling steps, and this number equals the index of the underlying matrix A, say ν. In the above example, the number of shuffling steps is two, so $\nu = 2$. Second, the rows shuffled in the first through kth shuffle steps, $k = 1,2,\ldots$, form a basis of the (row) null space $\{x: x^{\mathrm{T}}A^k = 0\}$. In the above example, $\{(\frac{1}{3}, \frac{1}{3}, \frac{2}{3}, 1)\}$ is a basis of $\{x: x^{\mathrm{T}}A = 0\}$, and $\{(\frac{1}{3}, \frac{1}{3}, \frac{2}{3}, 1), (-\frac{1}{3}, -\frac{1}{3}, -\frac{5}{3}, 0)\}$ is a basis of $\{x: x^{\mathrm{T}}A^2 = 0\}$. Finally, we show that A^D, the Drazin inverse of A, can be obtained from the matrix $\hat{A}$ on the right-hand side of the terminating matrix by computing $\hat{A}^{\nu+1}A^{\nu}$. In the above example

$$A^D = 9^{-3}\begin{pmatrix} 16 & -20 & -19 & 0 \\ 4 & -5 & -16 & 0 \\ -4 & 5 & 7 & 0 \\ -3 & 6 & 12 & 0 \end{pmatrix}^3 \begin{pmatrix} 2 & 4 & 6 & 5 \\ 1 & 4 & 5 & 4 \\ 0 & -1 & -1 & 0 \\ -1 & -2 & -3 & 3 \end{pmatrix}^2$$

$$= \begin{pmatrix} 3 & -1 & 2 & 2 \\ 2 & 1 & 3 & 3 \\ -1 & 0 & -1 & -1 \\ -1 & 0 & -1 & -1 \end{pmatrix}.$$

We next describe our algorithm in detail. Consider a given $n \times n$ real matrix A. In the course of the algorithm a sequence of pairs of matrices

$(A^{(k)}, B^{(k)})$, is generated, where $(A^{(0)}, B^{(0)}) = (A, I)$. Given $(A^{(k)}, B^{(k)})$ we execute row operations on $A^{(k)}$ to convert it into a matrix whose nonzero rows are linearly independent; moreover, if $A^{(k)}$ is found to be nonsingular the algorithm terminates. Simultaneously, we execute the same row operations on $B^{(k)}$. Let $\bar{A}^{(k)}$ and $\bar{B}^{(k)}$ be the result of executing the above row operations on $A^{(k)}$ and $B^{(k)}$, respectively. If $\bar{A}^{(k)}$ has zero rows, we exchange these rows with the corresponding rows of $\bar{B}^{(k)}$ and proceed to iteration $k+1$. We show that if ν is the index of A, then the algorithm will always terminate on exactly the νth iteration. Moreover, the rows shuffled on iterations $0, \ldots, k-1$, for $k = 1, \ldots, \nu$, are a basis of the left null space of A^k. In addition, we show that if on iteration ν, $A^{(\nu)}$ is transformed into the identity matrix, $\bar{A}^{(\nu)} = I$, and $\bar{A}$ is defined to be the resulting matrix $\bar{B}^{(\nu)}$, then the Drazin inverse of A is equal to $\hat{A}^{\nu+1}A^{\nu}$.

A representation of the Drazin inverse of matrices for which zero is a simple eigenvalue is given in [8, Lemma 5.1]. The representation in this case (for which the index is known to be one) reduces to the execution of our shuffle algorithm. A special case of the above representation for matrices having the form $I-P$, where P is an irreducible stochastic matrix, is given in [4, Theorem 8].

A survey of methods for computing the index and Drazin inverse of a matrix can be found in [2]. In particular, efficient methods for computing the Drazin inverse are given in [5] and [6, 7].

Further details and proofs concerning the new shuffle algorithm can be found in a forthcoming paper of the authors that will appear in this journal.

The authors wish to thank Eric V. Denardo for illuminating comments concerning the operation of the algorithm, and constructive criticism of the paper.

REFERENCES

1 K. M. Anstreicher, Generation of feasible descent directions in continuous time linear programming, Technical Report SOL 83-18, Dept. of Operations Research, Stanford Univ., Stanford, Calif., 1983.

2 S. L. Campbell and C. D. Meyer, *Generalized Inverses of Linear Transformations*, Fearon Pitman, Belmont, Calif., 1979.

3 S. L. Campbell, C. D. Meyer, and N. J. Rose, Applications of the Drazin inverse to linear systems of differential equations with singular constant coefficients, *SIAM. J. Appl. Math.* 31:411–425 (1976).

4 E. V. Denardo, A primer on Markov chains, unpublished manuscript, 1978.

5 T. N. E. Greville, The Souriau-Frame algorithm and the Drazin pseudoinverse, *Linear Algebra Appl.* 6:205–208 (1973).

6 R. E. Hartwig, More on the Souriau-Frame algorithm and the Drazin inverse, *SIAM J. Appl. Math.* 31:42–46 (1976).
7 R. E. Hartwig, A method for computing $A^{(l)}$, *Math. Japon.* 26:37–43 (1981).
8 M. Haviv and U. G. Rothblum, Bounds on distances between eigenvalues, *Linear Algebra Appl.* 63:101–118 (1984).
9 D. G. Luenberger, Time-invariant descriptor systems, *Automatica* 14:473–480 (1978).
10 G. Strang, *Linear Algebra and Its Applications*, 2nd ed., Academic, New York, 1980.
11 P. Van Dooren, The computation of Kronecker's canonical form of a singular pencil, *Linear Algebra Appl.* 27:103–140 (1979).
12 J. H. Wilkinson, Note on the practical significance of the Drazin inverse, in *Recent Applications of Generalized Inverses* (S. L. Campbell, Ed.), Fearon-Pitman, Belmont, Calif., 1982, pp. 82–99.
13 S. Winograd, *Arithmetic Complexity of Computations*, SIAM, Philadelphia, 1980.

ON BIHOLOMORPHIC AUTOMORPHISMS OF THE UNIT BALL OF UNITARY MATRIX SPACES

by JONATHAN ARAZY[3]

Let S denote the space of all complex matrices $a = (a_{i,j})_{i,j=1}^{\infty}$ with only finitely many nonzero entries. A norm α on S is *unitarily invariant* if

$$\alpha(uaw) = \alpha(a)$$

for all $a \in S$ and all unitary matrices u, w with the property that $u_{i,j} = w_{i,j} = \delta_{i,j}$ if $\max\{i, j\}$ is large enough. For normalization one requires also that $\alpha(a) = 1$ for every rank-one partial isometry a in S.

The unitary matrix space S_α associated with α is the completion of S under α (these spaces are called also "unitary ideals" or "symmetric normed ideals"; see [1]). We denote by B_α the open unit ball of S_α. It is known that the most general unitarily invariant norm α on S is given by

$$\alpha(a) = \beta(s_n(a)),$$

where $s_n(a) = \lambda_n((a^*a)^{1/2})$, $n = 1, 2, \ldots$, are the singular numbers of a, and

[3]Department of Mathematics, University of Haifa, Haifa 31999, Israel.

β is a symmetric norm on sequences. For $1 \leqslant p < \infty$ we put

$$\alpha_p(a) = \left(\sum_{n=1}^{\infty} s_n(a)^p \right)^{1/p} = \|(s_n(a))\|_{l_p}$$

and

$$\alpha_\infty(a) = s_1(a) = \text{the operator norm of } a.$$

We let $S_p = S_{\alpha_p}$. Thus S_1 is the trace class, S_2 is the space of Hilbert-Schmidt operators, and S_∞ is the space of compact operators on l_2.

It is well known that B_2 and B_∞ are bounded symmetric domains, that is the corresponding groups of biholomorphic automorphisms act transitively (in fact, by Möbius transformations; see [2] and [5]). If these two cases are ruled out, the situation changes drastically.

THEOREM 1. *Let α be a unitarily invariant norm, different from α_2 and α_∞. Then every biholomorphic automorphism of B_α extends to a linear isometry of S_α.*

With the aid of [3] we conclude

COROLLARY. *Let α be as in Theorem 1, and let F be a biholomorphic automorphism of B_α. Then there exist unitary matrices u, w so that either*

$$F(a) = uaw, \qquad a \in S_\alpha,$$

or

$$F(a) = ua^T w, \qquad a \in S_\alpha,$$

where a^T is the transpose of a.

A holomorphic vector field $X: B_\alpha \to S_\alpha$ is called *complete* if there exists a solution $\phi = \phi_X: R \times B_\alpha \to B_\alpha$ to the initial-value problem

$$\frac{\partial}{\partial t}\phi(t, a) = X(\phi(t, a)),$$

$$\phi(0, a) = a$$

for every $a \in B_\alpha$.

THEOREM 2. *Let α be a unitarily invariant norm, different from α_2 and α_∞. Then every complete holomorphic vector field $X: B_\alpha \to S_\alpha$ extends to a bounded, linear, skew-Hermitian operator on S_α.*

Here "skew-Hermitian" means that the numerical range is purely imaginary. Again, using [3] we get

COROLLARY. *Let α and $X: B_\alpha \to S_\alpha$ be as in Theorem 2. Then there exist bounded Hermitian matrices b, c so that*

$$X(a) = i(ba + ac), \qquad a \in S_\alpha.$$

The proofs of Theorems 1 and 2 use the "contraction principle" of [4].

REFERENCES

1 I. C. Gohberg and M. G. Krein, Introduction to the theory of linear non-selfadjoint operators, *Amer. Math. Soc. Transl.* 18 (1969).
2 L. A. Harris, Bounded symmetric homogeneous domains in infinite dimensional spaces, in *Lecture Notes in Math.*, Vol. 364, Springer, New York, 1973, pp. 13–40.
3 A. R. Sorour, Isometries of norm ideals of compact operators, *J. Funct. Anal.* 43:69–77 (1981).
4 L. L. Stachó, A projection principle concerning biholomorphic automorphisms, *Acta Sci. Math. (Szeged)* 44:99–124 (1982).
5 H. Upmeir, *Symmetric Banach Manifolds and Jordan C* -Algebras*, Mathematics Studies No. 104, North Holland, New York, 1985.

THE MAXIMUM DISTANCE PROBLEM AND BAND SEQUENCES

by A. BEN-ARTZI,[4*] R. L. ELLIS,[5] I. GOHBERG,[4] and D. LAY[5]

In this paper we solve the following problem. For $1 \leqslant j, k \leqslant n$ and $|j-k| \leqslant m$, let a_{jk} be a given complex number with $a_{kj} = \bar{a}_{jk}$. We wish to find linearly independent vectors $x_1, \ldots, x_n$ such that $\langle x_k, x_j \rangle = a_{jk}$ for $|j-k| \leqslant m$ and such that the distance from x_k to the linear span of $x_1, \ldots, x_{k-1}$ is maximal for $2 \leqslant k \leqslant n$. We construct and characterize all such sequences of vectors. Our solution leads naturally to the class of m-band

[4]School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel-Aviv University, Tel-Aviv, Israel.
[5]Department of Mathematics, University of Maryland, College Park, Md.

sequences of vectors in an inner-product space. We study these sequences and characterize their equivalence classes under unitary transformations. A precise formulation of the problem now follows.

Maximum-Distance Problem

For any vectors $x_1,\dots,x_n$ in C^n and for $2 \leqslant k \leqslant n$, let $\mathrm{sp}\langle x_1,\dots,x_k\rangle$ denote the subspace spanned by $x_1,\dots,x_k$, and let $\mathrm{dist}(x_k,\mathrm{sp}\langle x_1,\dots,x_{k-1}\rangle)$ denote the distance from x_k to $\mathrm{sp}\langle x_1,\dots,x_{k-1}\rangle$. Given $0 \leqslant m < n$ and a set $\{a_{jk}: |j-k| \leqslant m\}$ of complex numbers satisfying $a_{kj} = \bar{a}_{jk}$, we shall say that a sequence of vectors $\{x_k\}_{k=1}^n$ is admissible if it is linearly independent and if

$$\langle x_k, x_j\rangle = a_{jk} \qquad (|j-k| \leqslant m). \tag{1.1}$$

We define

$$d_k = \sup \mathrm{dist}(x_k,\mathrm{sp}\langle x_1,\dots,x_{k-1}\rangle) \qquad (2 \leqslant k \leqslant n),$$

where the supremum is taken over all admissible sequences of vectors in C^n. The maximum distance problem is to describe all admissible sequences $\{x_k\}_{k=1}^n$ such that

$$\mathrm{dist}(x_k,\mathrm{sp}\langle x_1,\dots,x_{k-1}\rangle) = d_k \qquad (2 \leqslant k \leqslant n).$$

Each such sequence is called a solution of the maximum-distance problem. This maximum-distance problem has close connections with maximum entropy in the mathematical theory of signal processing [2, 4, 5].

Band Sequences

The study of the maximum-distance problem leads naturally to the notion of a band sequence of vectors in C^n. If we apply the Gram-Schmidt process to a sequence of vectors $\{y_k\}_{k=1}^n$, we obtain a sequence $\{\omega_k\}_{k=1}^n$ of orthonormal vectors that is related to $\{y_k\}$ by a system of equations that may be written in the form

$$a_{1k}y_1 + \cdots + a_{kk}y_k = \omega_k \qquad (k = 1,\dots,n). \tag{1.2}$$

DEFINITION. Let m and n be integers with $0 \leqslant m < n$. We say that a linearly independent sequence $\{y_k\}_{k=1}^n$ is an m-band sequence if $a_{jk} = 0$ for $m < k \leqslant n$ and $1 \leqslant j < k - m$.

We shall say that two linearly independent sequences $\{y_k\}_{k=1}^n$ and $\{z_k\}_{k=1}^n$ in C^N are equivalent if there is an $N \times N$ unitary matrix U such that $z_k = Uy_k$ for $1 \leqslant k \leqslant n$.

We prove that the set of solutions to the maximum distance problem consists of one such equivalence class of m-band sequences.

Different characterizations of m-band sequences are given. Finally a maximum-volume problem is considered and shown to be equivalent to the maximum-distance problem.

REFERENCES

1 E. Asplund, Inverses of matrices $\{a_{ij}\}$ which satisfy $a_{ij} = 0$ for $j > i + p$, *Math. Scand.* 7:57–60 (1959).

2 J. Burg, Maximum entropy spectral analysis, Ph.D. Dissertation, Stanford Univ., Stanford, Calif., 1975.

3 C. Davis, W. M. Kahan, and H. F. Weinberger, Norm-preserving dilations and their applications to optimal error bounds, *SIAM J. Numer. Anal.* 19:445–469 (1982).

4 H. Dym and I. Gohberg, Extensions of band matrices with band inverses, *Linear Algebra Appl.* 36:1–24 (1981).

5 R. L. Ellis, I. Gohberg, and D. C. Lay, Band extensions, maximum entropy and the permanence principle, in *Maximum Entropy and Bayesian Methods in Applied Statistics* (J. Justice, Ed.), Cambridge U.P., 1986.

6 I. Gohberg and S. Goldberg, *Basic Operator Theory*, Birkhäuser, Boston, 1981.

SIGN PATTERNS OF MATRICES AND THEIR INVERSES

by MARC A. BERGER[6*] and ALEXANDER FELZENBAUM[6]

Given an $n \times n$ (± 1)-matrix $S = (s_{ij})$, any $n \times n$ real matrix $M = (m_{ij})$ for which $\operatorname{sgn}(m_{ij}) = s_{ij}$ $(\forall_{i,j})$ is said to be a *realization* of S. Two (± 1)-matrices S and T are said to be *paired* if they have realizations which are inverse to one another. Define a graph $\mathcal{G}_n$ whose matrices are all the $n \times n$ (± 1)-matrices, and whose edges connect all paired matrices. Following the "Open Questions" of Johnson, Leighton, and Robinson [1], we study this graph $\mathcal{G}_n$.

[6] Department of Pure Mathematics, The Weizmann Institute of Science, Rehovot, Israel.

In addition to a complete characterization of $\mathscr{G}_3$, our results include:

(i) $\delta(\mathscr{G}_n) \geqslant 2^{[n/2]}$ (for $n \geqslant 3$).

(ii) For any permutation π on $\{1,\ldots,n\}$, S and $F_\pi(S)$ are paired, where

$$F_\pi(S) = \{2I - [S(\pi)]^t S\}[S(\pi)]^t$$

and $S(\pi) = (q_{ij})$, where

$$q_{ij} = \begin{cases} s_{ij}, & j = \pi(j), \\ 0, & j \neq \pi(j) \end{cases}.$$

Our analyses lead to good bounds on the radius, diameter, and connectivity of $\mathscr{G}_n$. We also obtain results concerning the pairing of $(0, \pm 1)$-matrices.

REFERENCES

1 C. R. Johnson, T. L. Leighton, and H. A. Robinson, Sign patterns of inverse-positive matrices, *Linear Algebra Appl.* 24:75–83 (1979).

HOLDABILITY, IRREDUCIBILITY, AND M-MATRICES

by ABRAHAM BERMAN[7*] and RONALD J. STERN[8]

Let $A \in R^{n,n}$, $B \in R^{n,m}$, and let U denote the class of piecewise continuous R^m-valued functions. Consider the control system

$$\dot{x}(t) = Ax(t) + Bu(t) \qquad t \geqslant 0 \tag{L}$$

where $u \in U$. A nonempty set $S \subseteq R^n$ is *holdable* [with respect to (L)] if for any initial state $x_0 \in S$ there exists a control function $u \in U$ such that $x(t, x_0, u) \in S \ \forall t \geqslant 0$.

In this paper we study problems which have to do with holdability of the nonnegative orthant R^n_+ and related problems in matrix theory.

---

[7] Department of Mathematics, Technion–Israel Institute of Technology, Haifa 32000, Israel. Research supported by the Fund for Promotion of Research at the Technion.

[8] Department of Mathematics, Concordia University, Montreal, Quebec, H4B 1R6, Canada. Research supported by NSERC Grant No. 4641.

A state $x_0 \in R^n$ is *controllable* to a target set S if there exist $t \geq 0$ and $u \in U$ such that $x(t, x_0, u) \in S$. The set $X_{A,B}(S)$ of all points which are controllable to S is

$$X_{A,B}(S) = X_A(S) + Y_{A,B},$$

where

$$X_A(S) = \bigcup_{t \geq 0} e^{-tA}(S)$$

and

$$Y_{A,B} = \text{range}[B; AB; A^2B; \ldots; A^{n-1}B].$$

PROBLEM 1. Characterize the set $X_{A,B}(R_+^n)$ when R_+^n is holdable.

In other words, it is desired to find the set of initial states from which one can "hit and hold" the target R_+^n.

Consider the uncontrolled system

$$\dot{x}(t) = Ax(t), \qquad t \geq 0. \tag{L_0}$$

The nonnegative orthant is *positively invariant* [with respect to (L_0)] if

$$e^{tA}(R_+^n) \subseteq R_+^n \qquad \forall t \geq 0.$$

This is equivalent to A being *essentially nonnegative*, i.e., $A + \alpha I \geq 0$ for some α. The nonnegative orthant is *strictly positively invariant* if $e^{tA}(R_+^n \setminus \{0\}) \subseteq \text{int } R_+^n \ \forall t > 0$. This is equivalent to A being essentially nonnegative and irreducible. The following theorem is well known.

THEOREM 1 [4, 9]. *If R_+^n is positively invariant then*

$$\lambda_A := \max\{\text{Re}(\lambda);\ \lambda \in \sigma(A)\}$$

is an eigenvalue of A and has an associated nonnegative eigenvector. If R_+^n is strictly positively invariant, then λ_a is a simple eigenvalue of A which has an associated positive eigenvector, v_A. Furthermore, A has only one unit nonnegative eigenvector.

When R_+^n is positively invariant, $\overline{X_A(R_+^n)}$ is a polyhedral cone [1]. Formulas for $\overline{X_A(R_+^n)}$ are given in [1] and [6] under spectral assumptions on

A, and in [6] when $Ax \leqslant 0$ for some $x > 0$. The problem of determining the reachability set $X_A(R_+^n)$ from its closure is analyzed in [7]. When R_+^n is strictly positively invariant, $X_A(R_+^n)$ is given by

$$X_A(R_+^n) = \{0\} \cup \operatorname{int}\{\operatorname{span}(v_A) + \operatorname{range}(A - \lambda_A I)\}. \qquad (*)$$

Returning to Problem 1, we first observe that holdability of the nonnegative orthant is equivalent to the existence of $F \in R^{m,n}$ such that R_+^n is positively invariant with respect to the controlled differential system

$$\dot{x}(t) = (A + BF)x(t), \qquad t \geqslant 0.$$

Consider now the control system

$$\dot{x}(t) = (A + BF)x(t) + Bu(t), \qquad t \geqslant 0. \qquad (\tilde{L})$$

It is easy to see that an initial state is controllable to a given target under system (L) if and only if it is controllable to the target under ($\tilde{L}$) and that $Y_{A,B} = Y_{A+BF,B}$. Thus when R_+^n is holdable and $A + BF$ is essentially nonnegative,

$$X_{A,B}(R_+^n) = X_{A+BF}(R_+^n) + Y_{A,B}$$

and the results in [1], [6] and [7] could potentially be employed in order to find the reachability set $X_{A+BF}(R_+^n)$. The situation is simpler when $A + BF$ is also irreducible and (*) can be employed.

Thus the analysis of Problem 1 gives rise to

PROBLEM 2. Find, if possible, a linear feedback law $u(t) = Fx(t)$ such that

$$x_0 \geqslant 0, t \geqslant 0 \Rightarrow x(t, x_0, u) \geqslant 0$$

and

PROBLEM 3. Find, if possible, a linear feedback law $u(t) = Fx(t)$ such that

$$0 \neq x_0 \geqslant 0, t > 0 \Rightarrow x(t, x_0, u) > 0,$$

which are equivalent to

PROBLEM 2_M. Find, if possible, a matrix F such that $A + BF$ is essentially nonnegative

and

PROBLEM 3_M. Find, if possible, a matrix F such that $A + BF$ is essentially nonnegative and irreducible.

Consider the following family of n linear programming problems:

$$(P_i) \qquad \text{maximize} \quad \varepsilon$$

$$\text{subject to} \quad A_{*i}(i|\,) + B(i|\,)u - \varepsilon e \geq 0,$$

$$\varepsilon \geq 0.$$

Here the unknowns are the scalar ε and the m-vector u. A_{*i} denotes the ith column of A, $X(i|\,)$ denotes the matrix obtained from X by deleting its ith row, and e is a column full of ones of height $n-1$.

THEOREM 2. *If $[\varepsilon_i, u_i]$ is feasible in (P_i), $i = 1,\ldots,n$, then $A + BF$ is essentially nonnegative for $F = [u_1;\ldots;u_n]$. If (P_i) is infeasible for some i, then there exists no F such that $A + BF$ is essentially nonnegative.*

Let α be a subset of $\{1,\ldots,n\}\backslash\{i\}$, and let e_α be an n-vector of 0's and 1's whose kth entry is 1 if and only if $k \in \alpha$. Consider the linear program

$$(P_{i,\alpha}) \qquad \text{maximize} \quad \varepsilon$$

$$\text{subject to} \quad A_{*i}(i|\,) + B(i|\,)u - \varepsilon e_\alpha(i|\,) \geq 0.$$

$$\varepsilon \geq 0.$$

A set α is *maximal* with respect to i if it is a maximal set for which $(P_{i,\alpha})$ has a positive value. Suppose all the programs (P_i) are feasible, $i = 1,\ldots,n$, and let $\alpha_{i1},\ldots,\alpha_{it_i}$ be the maximal sets with respect to i, $i = 1,\ldots,n$. For $1 \leq s_i \leq t_i$, $i = 1,\ldots,n$, define a graph $G_{s_1,\ldots,s_n}$ as a directed graph having vertices $1,\ldots,n$ and an arc from i to j if and only if $j \in \alpha_{is_i}$.

THEOREM 3. *Suppose all* (P_i) *are feasible,* $i = 1, \ldots, n$. *If a graph* $G_{s_1, \ldots, s_n}$ *is strongly connected and* (ε_i, u_i), $\varepsilon_i > 0$, *is a feasible solution of* $(P_{i, \alpha_{i,s_i}})$, *then* $A + BF$ *is (essentially nonnegative and) irreducible, where* $F = [u_1; \ldots; u_n]$. *If none of the graphs* $G_{s_1, \ldots, s_n}$, $1 \leqslant s_i \leqslant t_i$, $i = 1, \ldots, n$, *is strongly connected, then no such F exists.*

Consider now the "stabilizability-holdability" problem

PROBLEM 4. Find, if possible, a linear feedback law $u(t) = Fx(t)$ such that for any $x_0 \geqslant 0$ the application of this law results in $x(t, x_0, u) \geqslant 0 \; \forall t \geqslant 0$ and $x(t, x_0, u) \to 0$ as $t \to \infty$.

Recall that M is a *nonsingular M-matrix* [2,5,8] if $-M$ is essentially nonnegative and all eigenvalues of M have a positive real part. Thus Problem 4 is equivalent to

PROBLEM 4_M. Find, if possible, a matrix F such that

$$-(A + BF) \text{ is a nonsingular } M\text{-matrix.} \qquad (**)$$

Let $X[\alpha]$ denote the principal submatrix of X built on indices in α.

THEOREM 4. *A matrix F satisfies* (**) *if and only if*

$$(A + BF)_{ij} \geqslant 0 \qquad \text{if} \quad 1 \leqslant i \neq j \leqslant n$$

and

$$(-1)^k \det((A + BF)[\{1, \ldots, k\}]) > 0 \qquad \text{for} \quad k = 1, \ldots, n.$$

In the case of scalar input (i.e., $m = 1$*) the last set of inequalities become*

$$(-1)^k (A[\{1, \ldots, k\}]) + f_{11} d_{k1} + \cdots + f_{1k} d_{kk} > 0 \qquad \text{for} \quad k = 1, \ldots, n,$$

where d_{ij} *denotes the determinant obtained from* $A[\{1, \ldots, k\}]$ *by replacing its jth column by the column B.*

Details, examples, and proofs are given in [3].

REFERENCES

1 A. Berman, M. Neumann, and R. J. Stern, Cone reachability for nondiagonable linear differential systems, *Linear Algebra Appl.*, to appear.
2 A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic, New York, 1979.
3 A. Berman and R. J. Stern, Linear feedback, irreducibility and *M*-matrices, Technion preprint No. MT-706.
4 L. Elsner, Monotonie und randspektrum bei vollstetigen operatoren, *Arch. Rational Mech. Anal.* 36:356–365 (1970).
5 M. Fiedler and V. Pták, On matrices with nonpositive off-diagonal elements and positive principal minors, *Czechoslovak Math. J.* 12:382–400 (1962).
6 M. Neumann and R. Stern, Cone reachability for linear differential systems, *Applicable Anal.*, to appear.
7 ____, Boundary results for positively invariant cones and their reachability cones, *Linear and Multilinear Algebra* 17:143–154 (1985).
8 A. M. Ostrowski, Über die determinanten mit überwiegender hauptdiagonale, *Comment. Math. Helv.* 10:69–96 (1937).
9 H. Schneider and M. Vidyassagar, Cross positive matrices, *SIAM J. Numer. Anal.* 7:508–519 (1970).

ON EIGENSTRUCTURE AND SIGNAL PROCESSING

by A. M. BRUCKSTEIN[9]

Many important problems in signal processing may be formulated as follows. Observations, in the form of M-vectors $\underline{\tau}_k$, are made in order to obtain information on a set of parameters $\{\vartheta_1, \vartheta_2, \ldots, \vartheta_D\}$. The parameters ϑ_i influence the observations through "signature vectors" $a(\vartheta_i)$ of length M, in the following way:

$$\underline{\tau}_k = \begin{bmatrix} a(\vartheta_1) & \cdots & a(\vartheta_d) \end{bmatrix} \begin{bmatrix} s_1^k \\ \vdots \\ s_D^k \end{bmatrix} + \begin{bmatrix} n_1^k \\ \vdots \\ n^k \end{bmatrix} = \mathbf{A}(\Theta)\underline{S}_k + \underline{N}_k, \qquad (1)$$

where the vectors $\underline{S}_k$ and $\underline{N}_k$ are randomly (and independently) chosen and have covariance matrices $\mathbf{S}_{(D\times D)}$ and $\sigma^2\mathbf{I}_{(M\times M)}$ respectively. By assumption, the D signature vectors come from a known "signature manifold" $\{a(\vartheta)\}$ parameterized by ϑ, and we observe a set $\{\underline{\tau}_k\}_{k=0,1,\ldots,K}$ of randomly weighted and noisy linear combinations of D particular signatures. Our aim

[9] Faculty of Electrical Engineering, Technion–I.I.T., Haifa, Israel.

is to estimate the number D, the parameter values $\{\vartheta_1, \vartheta_2, \ldots, \vartheta_D\}$, the noise variance σ^2, and the covariance matrix $\mathbf{S}$, from the observations $\{\underline{\tau}_k\}_{k=0,1,\ldots,K}$. From (1) it follows that the covariance of the observations has the form

$$\mathbf{R} = E\underline{\tau}_k\underline{\tau}_k^T = \mathbf{A}(\Theta)\mathbf{S}\mathbf{A}^T(\Theta) + \sigma^2\mathbf{I}. \tag{2}$$

If $D < M$, and $\{a(\vartheta)\}$ has the property that $\mathbf{A}(\Theta)$ has full rank for any choice of $\{\vartheta_1, \vartheta_2, \ldots, \vartheta_D\}$, an eigenstructure analysis of $\mathbf{R}$ provides a complete solution of the problem, via the so-called multiple-signal characterization (MUSIC) algorithm, due to Ralph Schmidt [1]. The algorithm exploits the fact that the minimal eigenvalue of $\mathbf{R}$ is σ^2, its multiplicity is $M - D$, and the $M - D$-dimensional subspace spanned by the eigenvectors corresponding to the minimal eigenvalue is orthogonal to the D-dimensional subspace spanned by the columns of $\mathbf{A}(\Theta)$. Examples of signal-processing problems that fit the above framework are multitarget tracking with sensor arrays, the extraction of sinusoids from noise, spectral estimation, radar echo location and several other signal resolution problems; see e.g. [1]–[3].

REFERENCES

1 R. O. Schmidt, Multiple emitter location and signal parameter estimation, in *Proceedings of the R.A.D.C. Spectrum Estimation Workshop*, Oct. 1979; also A signal subspace approach to multiple emitter location and spectral estimation, Ph.D. Dissertation, Stanford Univ., Stanford, Calif., Nov. 1981.

2 M. Wax, T.- J. Shan, and T. Kailath, Spatio-temporal spectral analysis by eigenstructure methods, *IEEE Trans. Acoust. Speech Signal Process.* ASSP-32:817–827 (1984).

3 A. M. Bruckstein, T.- J. Shan, and T. Kailath, The resolution of overlapping echos, *IEEE Trans. Acoust. Speech Signal Process.* ASSP-33:1357–1367 (1985).

ON LINEARLY CONSTRAINED ENTROPY MAXIMIZATION

by YAIR CENSOR[10]

1. Introduction

Linearly constrained entropy maximization is used in a variety of research fields. We mention three applications in matrix theory and describe some recent algorithmic developments.

[10] Department of Mathematics and Computer Science, University of Haifa, Mt. Carmel, Haifa 31999, Israel.

2. *Entropy Maximization and Matrix Theory*

Given $u \in \mathbb{R}^n_+$, $b \in \mathbb{R}^m$ and an $m \times n$ real matrix A where all columns, $a^i \in \mathbb{R}^n$, $i = 1,2,\ldots,m$, of A^T are nonzero, consider the entropy maximization problem

$$\text{Max}\left[-\sum_{j=1}^{n} x_j \log\left(\frac{x_j}{u_j}\right)\right] \quad \text{subject to} \quad Ax = b, \quad x \geqslant 0. \tag{1}$$

The need to solve this problem arises in various fields of applications, including transportation planning, statistics, geometric programming, image reconstruction from projections, image restoration, pattern recognition, and spectral analysis; see [5, 6, 10, 11, 12] for references.

In matrix theory and linear algebra, entropy maximization is encountered in several cases, including the following:

(i) *Permanents.* Minc's conjecture stating that if A is an $n \times n$ $(0,1)$-matrix with row sums $r_1,\ldots,r_n$, then $\text{per}(A) \leqslant \prod_{i=1}^{n}(r_i!)^{1/r_i}$ (see [14, Chapter 6.2]) was first proved by Bregman [2], by employing duality relations between (1) and its dual problem.

(ii) *Matrix balancing.* This is a procedure in which a diagonal matrix D is determined such that if $D^{-1}AD = [c^1,\ldots,c^n] = [r^1,\ldots,r^n]^T$ then $\|r^i\| = \|c^i\|$ for $i = 1,2,\ldots,n$, in an appropriate norm. This routine is recommended to precede the computation of eigenvalues of A; see [15]. It can be shown (see, e.g., [10, 12]) that the balancing problem is equivalent to entropy maximization.

(iii) *Matrix scaling.* Here diagonal matrices D_1 and D_2 are sought such that the scaled matrix D_1AD_2 has prescribed row and column sums; see [13]. This problem has applications in transportation planning, economics and statistics (see [12]) and is again equivalent to entropy maximization.

3. *Algorithms for Entropy Maximization*

Bregman's method [1], as studied and extended in [4] and [9], is a primal-dual convex programming algorithm of the row-action type [3]. It applies to a class of objective functions, called *Bregman functions* in [4], to which the negative of the $x \log x$ entropy function of (1) belongs. Bregman's method applies to linear equality-constrained problems as well as to linear inequality constraints. The main contribution in [4] is a storage-efficient extension of the method to linear interval constraints. In [9], relaxation parameters were introduced into the method in a geometrically insightful way.

The relaxed Bregman method for entropy maximization over linear *inequality* constraints is

ALGORITHM 1 (see [8, 9]).

Initialization: $z^0 \in \mathbb{R}_+^m$ is arbitrary, and $x^0 \in \mathbb{R}^n$ is such that $x_j^0 = \exp[(-A^T z^0)_j - 1]$ for $j = 1, \ldots, n$.

Iterative step: $x_j^{k+1} = x_j^k \exp(c_k a_j^{i(k)})$, $j = 1, \ldots, n$, where $c_k = \min(z_{i(k)}^k, \beta_k)$ and $z^{k+1} = z^k - c_k e^{i(k)}$. Here $e^{i(k)}$ is the $i(k)$th standard basis vector in $\mathbb{R}^m$, and β_k is the solution of

$$\sum_{j=1}^{n} a_j^{i(k)} x_j^k \exp\left(\beta_k a_j^{i(k)}\right) = \alpha_k b_{i(k)} + (1 - \alpha_k)\langle a^{i(k)}, x^k \rangle, \tag{2}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product and α_k are underrelaxation parameters which are bounded away from zero, i.e., $0 < \varepsilon \leqslant \alpha_k \leqslant 1$. The control sequence $\{i(k)\}$ is assumed to be almost cyclic.

The algorithm performs successive "entropy projections" onto the bounding hyperplanes of $Ax \leqslant b$ or onto certain hyperplanes parallel to them. It requires in each iterative step an inner-loop calculation to approximate β_k from Equation (2). When applying Algorithm 1 to (1), i.e., only equality constraints, the dual variables z^k are eliminated and $c_k = \beta_k$, for all k, in the iterative step.

Another algorithm for solving (1) is MART (multiplicative algebraic reconstruction technique); see [3, 12] for references. Its iterative step is

$$x_j^{k+1} = x_j^k \exp\left(M_k a_j^{i(k)}\right), \qquad j = 1, \ldots, n, \tag{3}$$

where

$$M_k = \lambda_k \log\left(\frac{b_{i(k)}}{\langle a^{i(k)}, x^k \rangle}\right), \tag{4}$$

$0 < \varepsilon \leqslant \lambda_k \leqslant 1$ and $\{i(k)\}$ almost cyclic. Conceptually, it has the advantage that M_k is given by a closed-form formula and no inner-loop iterations are needed as in Algorithm 1. If the constraint matrix A is a $(0,1)$-matrix, then MART coincides with Bregman's method. In [12], Lamond and Stewart identified various independently discovered entropy maximization algorithms as special cases of Bregman's method, but were unable to do so for MART in the non-$(0,1)$ case.

The relationship between MART and Bregman's algorithm is studied in [8], where it is shown that MART can be considered as a particular underrelaxation strategy in Bregman's algorithm. This discovery settles Lamond and Stewart's question and prompted the construction of a new hybrid algorithm which essentially replaces, in Algorithm 1, β_k by M_k of (4). The resulting new algorithm for entropy maximization over linear *inequalities* was proposed in [6] and proved in [8]. Recently we have extended the scope of Bregman's algorithm by showing [7] that it is applicable also to a function which is not a Bregman function in the sense of [4], namely the $\log x$ entropy known as Burg's entropy.

The author's work on entropy maximization described here was supported by NSF Grant ECS-8117908 and NIH Grant HL-28438 during several visits to the Medical Image Processing Group, Department of Radiology, Hospital of the University of Pennsylvania, Philadelphia, Pa. 19104 USA.

REFERENCES

1 L. M. Bregman, The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming, *U.S.S.R. Comput. Math. and Math. Phys.* 7(3):200–217 (1967).

2 L. M. Bregman, Some properties of nonnegative matrices and their permanents, *Dokl. Akad. Nauk SSSR* 211:27–30 (1973) [*Soviet Math. Dokl.* 14:945–949 (1973)].

3 Y. Censor, Row-action methods for huge and sparse systems and their applications, *SIAM Rev.* 23:444–466 (1981).

4 Y. Censor and A. Lent, An iterative row-action method for interval convex programming, *J. Optim. Theory Appl.* 34:321–353 (1981).

5 Y. Censor, Finite series expansions reconstruction methods, *Proc. IEEE* 71:409–419 (1983).

6 Y. Censor, T. Elfving, and G. T. Herman, Special-purpose algorithms for linearly constrained entropy maximization, in *Maximum Entropy and Bayesian Methods in Inverse Problems, II* (C. Ray Smith, Ed.), to appear; also available as Report #84, Mathematics Publication Series, Univ. of Haifa, Haifa, Israel, Aug. 1984.

7 Y. Censor and A. Lent, Optimization of "$\log x$" entropy over linear constraints, *SIAM J. Control Optim.*, to appear.

8 Y. Censor, A. R. De Pierro, T. Elfving, G. T. Herman, and A. N. Iusem, On some methods for entropy maximization, in preparation.

9 A. R. De Pierro and A. N. Iusem, A relaxed version of Bregman's method for convex programming, *J. Optim. Theory Appl.*, to appear.

10 T. Elfving, On some methods for entropy maximization and matrix scaling, *Linear Algebra Appl.* 34:321–339 (1980).

11 J. N. Kapur, Twenty-five years of maximum-entropy principle, *J. Math. Phys. Sci.* 17:103–156 (1983).
12 B. Lamond and N. F. Stewart, Bregman's balancing method, *Transportation Res. Part B* 15B:239–248 (1981).
13 A. W. Marshall and I. Olkin, Scaling of matrices to achieve specified row and column sums, *Numer. Math.* 12:83–90 (1968).
14 H. Minc, *Permanents*, Addison-Wesley, Reading, Mass., 1978.
15 B. N. Parlett and C. Reinsch, Balancing a matrix for calculation of eigenvalues and eigenvectors, *Numer. Math.* 13:293–304 (1969).

CLASS FUNCTIONS OF FINITE GROUPS, NONNEGATIVE MATRICES, AND GENERALIZED CIRCULANTS

by DAVID CHILLAG[11]

Introduction

Using the regular representation of the algebra $\mathrm{cf}(G)$ of the complex class functions of a finite group G, we assign to each class function θ a matrix $M(\theta)$. The correspondence $\theta \to M(\theta)$ enables us to deduce results (old and new) on group characters from known theorems on nonnegative matrices. From matrix-theory point of view, the set $M(G) = \{M(\theta) \mid \theta \in \mathrm{cf}(G)\}$ enjoys many properties of circulant matrices. Different choices of G yield different sets of matrices, including the circulants, block circulants with circulant blocks of all levels, and many more.

Notation

Our group-theoretical notation and notions are taken mainly from the first four chapters of [9], and those on matrices from [8] and Chapter 2 of [2]. Here is some additional notation.

Let G be a finite group, and $C_1, C_2, \ldots, C_k$ all the conjugacy classes of G with $C_1 = \{1\}$. Let $\alpha = (\chi_1, \chi_2, \ldots, \chi_k)$ be any enumeration of the set $\mathrm{Irr}(G)$ of all irreducible complex characters of G. If θ is a class function of G, the matrix $M^\alpha(\theta)$ is defined to be the $k \times k$ matrix $M^\alpha(\theta) = \left(m_{ij}^\alpha(\theta)\right)$, where $m_{ij}^\alpha(\theta) = [\theta\chi_i, \chi_j]$, $1 \leqslant i, j \leqslant k$; here $[\ ,\ ]$ means inner product. The set $M(G, \alpha) = \{M^\alpha(\theta) \mid \theta \in \mathrm{cf}(G)\}$ will be called a family of G-circulants; its elements are G-circulants. It is clear that if $M(G, \alpha)$ and $M(G, \alpha')$ are two

[11] Department of Mathematics, Technion–Israel Institute of Technology, Haifa 32000, Israel.

families of G-circulants, then $M(G, \alpha') = \{P^t M P \mid M \in M(G, \alpha)\}$ for some permutation matrix P.

Finally, let $X(G, \alpha) = (\chi_i(C_j))$, $1 \leq i, j \leq k$, be the character-table matrix of G.

Examples of G-circulants

In [3], [4], [10] finite groups have been used to generalize circulants. In most of these places G is assumed to be abelian, in which case $G \cong \mathrm{Irr}(G)$ and some of the generalized circulants are similar to ours. For nonabelian groups [3, 10] the set of matrices is different from ours. Here are some examples:

(1) If G is a cyclic group of order n, then for some α, $M(G, \alpha)$ is the set of $n \times n$ circulants and $X(G, \alpha) = \sqrt{n}\, F_n^*$ (see p. 32 of [8] for notation).

(2) If $G = C_{n_1} \times C_{n_2} \times \cdots \times C_{n_r}$ is a general abelian group where C_{n_i} is a cyclic group of order n_i, then for some $\alpha_1, \alpha_2, \ldots, \alpha_n, \alpha$ we get that $M(G, \alpha)$ is a level-r block circulant of type $(n_1, n_2, \ldots, n_r)$ and $X(G, \alpha) = (n_1 n_2 \cdots n_r)^{1/2} F_{n_1}^* \otimes F_{n_2}^* \otimes \cdots \otimes F_{n_r}^*$ (see p. 188 of [8] for definitions); here $\otimes$ means the Kronecker product.

(3) Example (2) is a special case of the general direct product $G = G_1 \times G_2 \times \cdots \times G_r$. Here for some α the family $M(G, \alpha)$ is a family of block G_1-circulants in which the blocks are block G_2-circulants in which the blocks are block G_3-circulants, etc. See [6] for the precise statement and further examples. Here also $X(G, \alpha) = X(G_1, \alpha_1) \otimes X(G_2, \alpha_2) \otimes X(G_r, \alpha_r)$ for some $\alpha_1, \alpha_2, \ldots, \alpha_r$. This example enables us to construct new examples from old ones by forming block matrices of all levels. See [6] for details.

Properties of G-circulants

These are summed up in the next theorem. In view of the above examples, all the following results from [8] are special cases of this theorem: Theorem 3.1.1 and its corollary; all the results of pp. 68, 72 and 73; Theorems 3.2.4, 3.3.1, 3.4.20, 3.4.21, 5.8.1, 5.8.2, 5.8.3, 5.8.4, 5.8.5, and others. For details and proofs see [6].

THEOREM 1. *Let G, k, C_i, χ_i, α, and $X = X(G, \alpha)$ be as defined above. Then:*

(a) *The family $M(G, \alpha)$ of G-circulants is a commutative algebra over the complex number field, and the mapping $\theta \to M^\alpha(\theta)$ is an algebra isomorphism from $\mathrm{cf}(G)$ onto $M(G, \alpha)$.*

(b) *$M(G, \alpha)$ is a set of commutative normal matrices which are simultaneously diagonalized by X. In fact $X^{-1} M^\alpha(\theta) X = \mathrm{diag}(\theta(C_1)$,*

$\theta(C_2), \ldots, \theta(C_k))$ *for each* $\theta \in \mathrm{cf}(G)$. *Also* $XDX^* = I$, *where* $D = (1/|G|)\mathrm{diag}(|C_1|, |C_2|, \ldots, |C_k|)$.

(c) *The set* $\{M^\alpha(\chi_i) \mid 1 \leqslant i \leqslant 1\}$ *is a basis of* $M(G, \alpha)$. *Furthermore, if* $M = \sum_{i=1}^k \alpha_i M^\alpha(\chi_i)$ *is an arbitrary element of* $M(G, \alpha)$, *then the eigenvalues of* M *are* $\sum_{i=1}^k \alpha_i \chi_i(C_j)$, $j = 1, 2, \ldots, k$. *In particular the eigenvalues of* $M^\alpha(\chi_i)$ *are the entries of the ith row of* X.

(d) *If* A *is any* $k \times k$ *complex matrix, then the following statements are equivalent*: (i) $A \in M(G, \alpha)$; (ii) A *commutes with* $M^\alpha(\chi_i)$ *for all* $1 \leqslant i \leqslant k$; (iii) $A = X \cdot \mathrm{diag}(\tau_1, \tau_2, \ldots, \tau_k) \cdot X^{-1}$ *for some complex numbers* $\tau_1, \tau_2, \ldots, \tau_k$.

Characters and Nonnegative Matrices

If θ is a character of G, then $M^\alpha(\theta)$ has further properties.

THEOREM 2. *Using the notation of Theorem* 1, *let* θ *be a character of* G. *Then*:

(a) $M^\alpha(\theta)$ *is a nonnegative-integer matrix. Its leading (Perron-Frobenius) eigenvalue is* $\theta(1)$ *with a corresponding eigenvector* $(\chi_1(1), \chi_2(1), \ldots, \chi_k(1))^t$.

(b) $M^\alpha(\theta)$ *is irreducible if and only if* θ *is faithful.*

(c) $M^\alpha(\theta)$ *is primitive if and only if* $Z(\theta) = 1$, *in which case* $\gamma(M^\alpha(\theta)) = \mathrm{ccn}(\theta)$ (*see* [2] *and* [1] *for definitions*).

(d) *Let* $\alpha_1, \alpha_2, \ldots, \alpha_m$ *be all the distinct values of* θ. *Then the characteristic polynomial* $\prod_{i=1}^k [x - \theta(C_i)]$ *and the minimal polynomial* $\prod_{i=1}^k (x - \alpha_i)$ *of* $M^\alpha(\theta)$ *have integer coefficients.*

The proof can be found in [5, 7]. Each of these properties can be applied to obtain results on characters (old and new) from those on matrices. See [1], [5], and [7] for such applications.

REFERENCES

1 Z. Arad, D. Chillag, and M. Herzog, Powers of characters of finite groups, in *J. Algebra*, to appear.

2 A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic, 1979.

3 R. Chalkley, Information about group matrices, *Linear Algebra Appl.* 38:121–133 (1981).

4 R. Chalkley, Matrices derived from finite abelian groups, *Math. Mag.* 49:121–129 (1976).

5 D. Chillag, Nonnegative matrices and products of characters and conjugacy classes in finite groups, *Publicationes Math. Debrecen.*, to appear.

6 D. Chillag, Generalized circulants and class functions of finite groups, Technion preprint No. MT-714.
7 D. Chillag, Character values of finite groups as eigenvalues of nonnegative integer matrices, *Proc. Amer. Math. Soc.*, to appear.
8 P. J. Davis, *Circulant Matrices*, Wiley, New York, 1979.
9 I. M. Isaacs, *Character Theory of Finite Groups*, Academic, 1976.
10 K. Wang, On the generalizations of circulants, *Linear Algebra Appl.* 25:197–218 (1979); Resultants and group matrices, *ibid.* 33:111–122 (1980); On the group matrices for a generalized group, *ibid.* 39:83–89 (1981); On the generalization of a retrocirculant, *Linear Algebra and Appl.*, *ibid.* 37:35–43 (1981).

ON THE SPECTRUM OF THE MATRIX $|x_i - x_j|^\alpha$

by NIRA DYN,[12*] TIMOTHY GOODMAN,[13] and CHARLES A. MICCHELLI[14]

In this work we investigate the spectrum of matrices of the form $A_n^\alpha = \{|x_i - x_j|^\alpha\}_{i,j=1}^n$, with $\alpha > 0$ and distinct $x_1, \ldots, x_n \in R$. These matrices occur in the solution of the interpolation problem by functions from $\text{span}\{|x - x_i|^\alpha,\ i = 1, \ldots, n\}$, for data points at $x_1, \ldots, x_n$. For α an odd integer, the above interpolating space is contained in the space of splines of degree α with simple knots $x_1, \ldots, x_n$. Similar matrices occur in the problem of multivariate interpolation by the radial functions $\{\|x - x^i\|_2^\alpha,\ i = 1, \ldots, n\}$ at the data points $x^1, \ldots, x^n \in R^s$ [2, 3]. Our analysis applies to the univariate situation only, and guarantees the nonsingularity of these matrices for any choice of distinct points $x_1, \ldots, x_n$. The case $0 < \alpha < 2$ is treated in [3] in the context of points in R^s. It is shown that A_n^α has $n - 1$ negative eigenvalues and one positive eigenvalue. In [1] it is shown that for $\alpha = 3$ and distinct $x_1, \ldots, x_n \in R$, A_n^α has $n - 2$ positive eigenvalues and 2 negative eigenvalues. Here we extend these results to

THEOREM 1. *Let $x_1, \ldots, x_n \in R$ be distinct, and $2l - 2 < \alpha < 2l$, $l \geq 1$. Then the matrix $B_n^\alpha = (-1)^l\{|x_i - x_j|^\alpha\}_{i,j=1}^n$ with $n \geq 2l - 1$ has $n - l$ positive eigenvalues and l negative eigenvalues.*

THEOREM 2. *Let $x_1, \ldots, x_n \in R$ be distinct, and let $\alpha \geq n > 0$. Then the matrix $A_n^\alpha = \{|x_i - x_j|^\alpha\}_{i,j=1}^n$ has $n/2$ positive eigenvalues and $n/2$ nega-*

[12]School of Mathematical Sciences, Tel-Aviv University, Tel-Aviv, Israel.
[13]Department of Mathematics, University of Dundee, Dundee, Scotland.
[14]IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598.

tive eigenvalues if n is even, while if n is odd, $(-1)^{[n/2]}A_n^\alpha$ *has* $[n/2]+1$ *positive eigenvalues and* $[n/2]$ *negative eigenvalues.*

THEOREM 3. *For* $\alpha > 0$ *an even integer,* $\alpha = 2p$, *and* $n > \alpha$ *distinct* $x_1, \ldots, x_n \in R$, *the matrix* $(-1)^P B_n^\alpha$ *has* $p+1$ *positive eigenvalues,* p *negative eigenvalues, and* $n - 2p - 1$ *zero eigenvalues.*

These results are based on a basic property of the matrices A_n^α, $\alpha > 0$, which is proved in [3] in the context of points $x^1, \ldots, x^n$ in R^s. Here we cite the relevant result for R:

THEOREM 4 [3]. *For* $2l - 2 < \alpha < 2l$, $l \geqslant 1$, *and distinct* $x_1, \ldots, x_n \in R$, $n \geqslant l$, *the matrix* A_n^α *satisfies*

$$(-1)^l y^T A_n^\alpha y > 0$$

for all $y^T = (y_1, \ldots, y_n)$ *with the property*

$$y \neq 0, \quad \sum_{i=1}^{n} y_i x_i^k = 0, \qquad k = 0, 1, \ldots, l-1.$$

This result guarantees that B_n^α has at least $n - l$ positive eigenvalues. The proof that the remaining l eigenvalues are all negative requires information on the number of zeros of functions spanned by $\{|x - x_i|^\alpha,\ i = 1, \ldots, n\}$ for $3 \leqslant n \leqslant \alpha + 1$. The result is an extension of a result of Sylvester concerning zeros of polynomials [4, p. 408].

THEOREM 5. *Let* $f(x) = \sum_{i=1}^n c_i |x - x_i|^\alpha$, *where n is odd,* $3 \leqslant n \leqslant \alpha + 1$, *and where* $x_1, \ldots, x_n$ *are distinct points in R, and* $c_1, \ldots, c_n \in R$, $\sum_{i=1}^n c_i^2 > 0$. *Then the number of distinct zeros of f does not exceed* $n - 1$.

Theorem 5 together with Theorem 2 yields the positivity of all the minors of odd order p, $3 \leqslant p \leqslant \alpha + 1$, of the matrix $\{|y_i - x_j|^\alpha\}_{i,j=1}^n$, from which the following oscillation properties of the eigenvectors of A_n^α are concluded:

THEOREM 6. *Let* $\lambda_1, \ldots, \lambda_n$, $|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_n|$, *be the eigenvalues of* A_n^α, *with corresponding eigenvectors* $\mathbf{u}^1, \ldots, \mathbf{u}^n$, *and let m be an odd*

integer, $m \leqslant \min\{n, \alpha+1\}$. *Then*

$$\lambda_1 > 0, \qquad |\lambda_{2i-1}| > |\lambda_{2i}| \geqslant |\lambda_{2i+1}|, \ \lambda_{2i}\lambda_{2i+1} > 0, \quad i = 1, \ldots, \left[\frac{m}{2}\right].$$

Moreover, for any p *even and* q *odd,* $p, q \leqslant m$,

$$p - 1 \leqslant S^-\left[\sum_{i=p}^{m} \alpha_i \mathbf{u}^i\right], \qquad S^+\left[\sum_{i=1}^{q} \beta_i \mathbf{u}^i\right] \leqslant q - 1$$

whenever $\alpha_p^2 + \cdots + \alpha_m^2 \neq 0$, $\beta_1^2 + \cdots + \beta_q^2 \neq 0$. *In particular* $\mathbf{u}^1 > 0$, $2i - 1 \leqslant S^-(\mathbf{u}^{2i+l}) \leqslant S^+(\mathbf{u}^{2i+l}) \leqslant 2i$, $l = 0, 1$, $i = 1, \ldots, [m/2]$. *Here* $S^-(\mathbf{x})$ *and* $S^+(\mathbf{x})$ *denote the maximal numbers of sign changes in the components of the vector* $\mathbf{x}$ *when zero components are ignored, and when zero components are given arbitrary signs, respectively.*

The proofs of these results will be published in the Proceedings of the Netherlands Academy of Sciences Series A.

REFERENCES

1 L. P. Bos and K. Salkauskas, On the matrix $[|x_i - x_j|^3]$ and the cubic spline continuity equation, *J. Approx. Theory*, to appear.
2 R. Franke, Scattered data interpolation: Test of some methods, *Math. Comp.* 38:181–200 (1982).
3 C. A. Micchelli, Interpolation of scattered data: Distance matrices and conditionally positive definite functions *Constr. Approx.*, 2:11–22 (1986).
4 J. J. Sylvester, *Collected Mathematical Papers*, Vol. II, Cambridge, 1908.

POSITIVE RECIPROCAL MATRICES

by L. ELSNER[15] and V. MEHRMANN[15*]

1. Introduction

An $n \times n$ matrix $A = [a_{ij}]$ is called positive reciprocal if $a_{ij} > 0$ and $a_{ij} = 1/a_{ji}$ for all $i, j = 1, \ldots, n$.

Positive reciprocal matrices play an important role in many applications. Recently, in a method called the analytic hierarchy process, they have

[15] Fakultät für Mathematik, Universität Bielefeld, Postfach 8640, 4800 Bielefeld 1, Federal Republic of Germany.

become the major mathematical tools for solving decision problems in economics and politics (Saaty [5]). Consider the following example: A consumer is asked to give a priority list for n products. By comparing product x_i with x_j pairwise he gives numbers a_{ij}, meaning that he likes product x_i a_{ij} times as much as product x_j. Clearly the matrix $[a_{ij}]$ will be positive reciprocal, but unfortunately this method usually leads to inconsistent answers, i.e., if the consumer prefers x_j over x_i by a factor a_{ji} and x_i over x_k by a factor a_{ik}, then in general the factor a_{jk} for comparing x_j and x_k is not equal to $a_{ji}a_{ik}$. Thus, for practical methods, the matrix A is usually approximated by a matrix $B = [b_{ij}]$ which is consistent, i.e. satisfies $b_{jk} = b_{ji}b_{ik}$ for all $i, j, k \in \{1, \ldots, n\}$. An analysis of the different possibilities for approximating A is given in [5] and [3]. Clearly, every consistent matrix can be written as $B = [b_{ij}]$ with $b_{ij} = [w_i/w_j]$ for some vector $w = [w_1, \ldots, w_n]^T$. Typical choices for W are the left or right Perron eigenvectors of A or the vectors of geometric row means or geometric column means of A.

In [3] Fichtner asked under which conditions on the matrix A the right (left) Perron vector is equal to the vector of geometric row (column) means, respectively. In this paper we shall consider this question and also analyze the relationship of the class of positive reciprocal matrices to other important subclasses of the positive (nonnegative) matrices.

2. *Notation and Preliminaries*

We start with some necessary definitions:

DEFINITION 1. An $n \times n$ matrix $A = [a_{ij}]$ is called *positive reciprocal* if $a_{ij} > 0$ and $a_{ij} = 1/a_{ji}$ for all $i, j \in \{1, \ldots, n\}$. The class of $n \times n$ positive reciprocal matrices is denoted by $\mathrm{PR}_{\langle n \rangle}$.

An $n \times n$ matrix A is called *consistent* if $DAD^{-1} = ee^T$ for some nonsingular diagonal matrix D. [Here $e = (1, \ldots, 1)^T$.]

We have the following simple result (e.g. [5]).

THEOREM 1. *Let $A \in \mathrm{PR}_{\langle n \rangle}$. Then $\rho(A) \geqslant n$ and $\rho(A) = n$ if and only if A is consistent, where $\rho(A)$ denotes the spectral radius of A.*

DEFINITION 2. An $n \times n$ matrix $A = [a_{ij}]$ is called an *inverse M-matrix* if $a_{ij} \geqslant 0$ for all $i, j \in \{1, \ldots, n\}$, $B = [b_{ij}] = A^{-1}$ exists, and $b_{ij} \leqslant 0$ for all $i, j \in \{1, \ldots, n\}$ $i \neq j$.

An $n \times n$ matrix A is called *totally nonnegative* (*totally positive*) if every minor of A is nonnegative (positive).

DEFINITION 3. A matrix of the form

$$\begin{bmatrix} c_0 & c_1 & c_2 & \cdot & \cdot & \cdot & c_{n-1} \\ c_{-1} & c_0 & c_1 & \cdot & \cdot & \cdot & c_{n-2} \\ \cdot & & \cdot & \cdot & & & \cdot \\ \cdot & & & \cdot & \cdot & & \cdot \\ \cdot & & & & \cdot & \cdot & \cdot \\ \cdot & & & & & \cdot & c_1 \\ c_{1-n} & \cdot & \cdot & \cdot & \cdot & c_{-1} & c_0 \end{bmatrix}$$

is called a *Toeplitz* matrix. A Toeplitz matrix with $c_{-j} = c_{n-j}$ for all $j \in \{1, \ldots, n-1\}$ is called a *circulant* matrix.

3. *Main Results*

We have the following results on the relationship of $\mathrm{PR}_{\langle n \rangle}$ to other subclasses of nonnegative matrices:

PROPOSITION 3.1. *Let* $A \in \mathrm{PR}_{\langle n \rangle}$, $n > 1$. *Then* A *is not an inverse M-matrix.*

PROPOSITION 3.2. *Let* $A \in \mathrm{PR}_{\langle n \rangle}$, $n > 1$. *Then* A *is not a totally positive matrix.*

PROPOSITION 3.3. *Let* $A \in \mathrm{PR}_{\langle n \rangle}$. *Then* A *is a totally nonnegative matrix if and only if* A *is consistent.*

Now we consider the following problem posed by Fichtner [3]: Characterize all matrices in $\mathrm{PR}_{\langle n \rangle}$ with the property that the right (left) Perron vector is the vector of geometric row (column) means respectively. In other words, let $u = (u_1, \ldots, u_n)^T$, $v = (v_1, \ldots, v_n)^T$ be such that

$$v_i = \sqrt[n]{\prod_{j=1}^{n} a_{ji}}, \quad u_i = \sqrt[n]{\prod_{j=1}^{n} a_{ij}}, \qquad i = 1, \ldots, n; \tag{3.4}$$

then

$$Au = \rho(A)u \quad \text{and} \quad v^T A = \rho(A)v^T. \tag{3.5}$$

We denote the subclass of $\mathrm{PR}_{\langle n \rangle}$ of matrices satisfying (3.4), (3.5) by $\mathrm{PR}^c_{\langle n \rangle}$.

LEMMA 3.6. *Let* $A \in \mathrm{PR}^e_{\langle n \rangle}$. *Then* $PDAD^{-1}P^T \in \mathrm{PR}^e_{\langle n \rangle}$ *for all* $D = \mathrm{diag}(d_1, \ldots, d_n)$ *with* $d_i > 0$ *for all* $i \in \{1, \ldots, n\}$ *and* $\prod_{i=1}^n d_i = 1$, *and all permutation matrices* P.

In view of Lemma 3.6 it suffices to characterize all those $A \in \mathrm{PR}_{\langle n \rangle}$ having row products (column products) equal to 1 and all row sums and column sums equal.

THEOREM 3.7.

(a) *Let* $A \in \mathrm{PR}_{\langle n \rangle}$ *be a circulant matrix. Let* $D = \mathrm{diag}(d_1, \ldots, d_n)$ *with* $d_i > 0$, $i = 1, \ldots, n$, *and* $\prod_{i=1}^n d_i = 1$, *and let* P *be a permutation matrix. Then* $PDAD^{-1}P^T \in \mathrm{PR}^e_{\langle n \rangle}$.

(b) *Let* $n \leqslant 5$. *Then* $A \in \mathrm{PR}^e_{\langle n \rangle}$ *if and only if there exist a diagonal matrix* D *as in* (a) *and a permutation matrix* P *such that* $PDAD^{-1}P^T$ *is a circulant matrix.*

Counting the number of free parameters and the number of equations that characterize matrices in $\mathrm{PR}^e_{\langle n \rangle}$, one sees that one cannot expect that for large n the circulant matrices are the only matrices in $\mathrm{PR}^e_{\langle n \rangle}$ up to diagonal and permutational similarity. However, if we restrict ourselves to Toeplitz matrices, the parameter count is the same for both sets.

CONJECTURE 3.8. Let $A \in \mathrm{PR}_{\langle n \rangle}$ be a Toeplitz matrix. Then the following are equivalent:

(i) $A \in \mathrm{PR}^e_{\langle n \rangle}$.

(ii) There exist a permutation P and a diagonal matrix D with $d_i > 0$ for $i = 1, \ldots, n$ and $\prod_{i=1}^n d_i = 1$ such that $PDAD^{-1}P^T$ is a circulant matrix.

(iii) A is a circulant matrix.

REFERENCES

1 A. Berman and R. J. Plemmons, *Nonnegative Matrices in The Mathematical Sciences*, Academic, New York, 1979.

2 P. J. Davis, *Circulant Matrices*, Wiley, New York, 1979.

3 J. Fichtner, Some thoughts about the mathematics of the analytic hierarchy process, Bericht Nr. 8403, Hochschule der Bundeswehr, Neubiberg.

4 C. R. Johnson, T. Wang, and W. Beine, A note on right-left asymmetry in an eigenvector-ranking scheme, *J. Math. Psych.*, Jan. 1979, pp. 61–64.

5 T. L. Saaty, *The Analytic Hierarchy Process*, McGraw-Hill, New York, 1980.

6 L. G. Vargas, Analysis of sensitivity of reciprocal matrices, *Appl. Math. and Comput.* 12:301–320 (1983).

PARTIAL REALIZATIONS, GENERALIZED BEZOUTIANS, AND BLOCK-HANKEL-MATRIX INVERSION

by PAUL A. FUHRMANN[16]

This note summarizes results on block-Hankel-matrix inversion. The detailed results will appear separately. It extends work of the author in the scalar case [7, 8].

We consider the problem of inverting the block Hankel matrix

$$H = \begin{pmatrix} G_1 & \cdots & G_n \\ \vdots & & \vdots \\ G_n & \cdots & G_{2n-1} \end{pmatrix}.$$

We assume that this matrix is nonsingular. In particular this implies the G_i are square, say $p \times p$, matrices. The inversion problem is related to the partial-realization problem studied in system theory. In this connection refer to Kalman [10], Gragg and Lindquist [9]. We seek rational extensions

$$G_X(z) = \frac{G_1}{z} + \cdots + \frac{G_{2n-1}}{z^{2n-1}} + \frac{X}{z^{2n}} + \cdots$$

of minimal McMillan degree whose first $2n-1$ coefficients coincide with $G_1, \ldots, G_{2n-1}$. Such an extension is uniquely determined by the choice of $X = G_{2n}$.

By the nonsingularity of H there exist coprime matrix-fraction representations of the forms

$$G_X(z) = P_X(z)Q_X(z)^{-1} = \bar{Q}_X(z)^{-1}\bar{P}_X(z)$$

such that both Q_X and $\bar{Q}_X$ are monic.

THEOREM 1. *Under the previous assumptions there exist uniquely determined polynomial matrices A, B, $\bar{A}$, and $\bar{B}$ such that $Q_X^{-1}\bar{A}$ and $A\bar{Q}_X^{-1}$ are strictly proper and*

$$\begin{pmatrix} A & B \\ -\bar{Q}_X & \bar{P}_X \end{pmatrix}\begin{pmatrix} P_X & -\bar{B} \\ Q_X & \bar{A} \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}.$$

[16] Department of Mathematics, Ben-Gurion University of the Negev, Beer Sheva, Israel

In particular we have the two interesting relations $\bar{P}_X Q_X = \bar{Q}_X P_X$ and $\overline{AQ}_X = Q_X A$. In terms of polynomial models [3–6], we have two maps

$$Z: X_{Q_X} \to X_{\bar{Q}_X} \text{ and } \bar{Z}: X_{\bar{Q}_X} \to X_{Q_X}$$

defined by

$$Zf = \pi_{\bar{Q}_X} \bar{P}_X f \qquad \text{for} \quad f \text{ in } X_{Q_X}$$

and

$$\bar{Z}g = \pi_{Q_X} \bar{A}g \qquad \text{for} \quad g \text{ in } X_{\bar{Q}_X}.$$

By the Bezout equations derivable from the matrix identity, the maps Z and $\bar{Z}$ are inverses of each other. Moreover they are very closely related to the inversion problem.

The polynomial matrices in the matrix identity can be used, following Anderson and Jury [2], to define the generalized Bezoutian $\Gamma(\bar{A}, \bar{Q}_X, Q_X, A) = (B_{ij})$ by

$$\frac{\bar{A}(z)\bar{Q}_X(w) - Q_X(z)A(w)}{z - w} = \sum\sum B_{ij} z^{i-1} w^{j-1}.$$

The generalized Bezoutian, in conjunction with polynomial model theory, can be used to get the following extension of a result of Lander [11].

THEOREM 2. *Under the assumptions of invertibility of H, then H^{-1} can be obtained from a minimal rational extension of the sequence $G_1, \ldots, G_{2n-1}$ by $H^{-1} = \Gamma(\bar{A}, \bar{Q}_X, Q_X, A)$, where A, B, $\bar{A}$, and $\bar{B}$ are the polynomial matrices determined by the matrix equality.*

We can parametrize the set of all minimal rational extensions. This is a special case of a more general formula of Antoulas [1].

THEOREM 3. *Given the nonsingular Hankel matrix H. Let $G = PQ^{-1} = \bar{Q}^{-1}\bar{P}$ and $G_X = P_X Q_X^{-1} = \bar{Q}_X^{-1}\bar{P}_X$ be the minimal rational extensions of $G_1, \ldots, G_{2n-1}$, corresponding to the choices $G_{2n} = 0$ and $G_{2n} = X$ respectively, and all matrix fractions taken coprime with monic denominators. Let A, B, $\bar{A}$, and $\bar{B}$ be determined as before. Then all minimal rational extensions are parametrized by*

$$G_X = (P + \bar{B}X)(Q - \bar{A}X)^{-1} = (\bar{Q} - XA)^{-1}(\bar{P} + XB).$$

REFERENCES

1 A. C. Antoulas, On recursiveness and related topics in linear systems, Technical Report TR 8507, Rice Univ., 1985.
2 B. D. O. Anderson and E. I. Jury, Generalized Bezoutian and Sylvester matrices in multivariable linear control, *IEEE Trans. Automat. Control.* AC-21:551–556 (1976).
3 P. A. Fuhrmann, Algebraic system theory: An analyst's point of view, *J. Franklin Inst.* 301:521–540 (1976).
4 P. A. Fuhrmann, Linear feedback via polynomial models, *Internat. J. Control* 30:363–377 (1979).
5 P. A. Fuhrmann, Duality in polynomial models with some applications to geometric control theory, *IEEE Trans. Automat. Control* AC-26:284–295 (1981).
6 P. A. Fuhrmann, On symmetric rational transfer functions, *Linear Algebra Appl.* 50:167–250 (1983).
7 P. A. Fuhrmann, On the partial realization problem and the recursive inversion of Hankel and Toeplitz matrices, in *Contemp. Math. Linear Algebra and its Role in System Theory*, Amer. Math. Soc., 1985.
8 P. A. Fuhrmann, Remarks on the inversion of Hankel matrices, *Linear Algebra Appl.*, to appear.
9 W. B. Gragg and A. Lindquist, "On the partial realization problem," *Linear Algebra Appl.* 50:277–319 (1983).
10 R. E. Kalman, On partial realizations, transfer functions and canonical forms, *Acta Polytech. Scand.* 31:9–32 (1979).
11 F. I. Lander, The Bezoutian and inversion of Hankel and Toeplitz matrices (in Russian), *Mat. Issled.* IX(32):69–87 (1974).

MINIMALITY AND IRREDUCIBILITY OF TIME-INVARIANT LINEAR BOUNDARY-VALUE SYSTEMS

by I. GOHBERG,[17] M. A. KAASHOEK,[18] and L. LERER[19*]

This presentation concerns the canonical structure of linear time-invariant systems with well-posed boundary conditions (briefly: boundary-value systems). In state-space representation such a system has the form

$$\theta:\begin{cases} x(t) = Ax(t) + Bu(t), & a \leqslant t \leqslant b, \\ y(t) = Cx(t), & a \leqslant t \leqslant b, \\ N_1 x(a) + N_2 x(b) = 0, \end{cases}$$

[17] School of Mathematical Science, Tel-Aviv University, Tel-Aviv, Israel.
[18] Department of Mathematics, Vrife Universiteit, The Netherlands.
[19] Department of Mathematics, Technion–I.I.T., Haifa 32000, Israel.

where $A: X \to X$, $B: Z \to X$ and $C: X \to Y$ are linear operators acting between finite-dimensional inner-product spaces. The boundary conditions of θ, defined by the linear operators N_1 and N_2 acting on the *state space* X, are assumed to be well posed, i.e., $\det(N_1 e^{aA} + N_2 e^{bA}) \neq 0$. In this case θ has a well-defined *input-output map*, namely the integral operator

$$T_\theta: L_2([a,b], Z) \to L_2([a,b], Y),$$

$$y(t) = (T_\theta u)(t) = \int_a^b k_\theta(t,s)u(s)\,ds, \qquad a \leqslant t \leqslant b,$$

of which the kernel k_θ is given by

$$k_\theta(t,s) = \begin{cases} Ce^{(t-a)A}(I-P)e^{-(s-a)A}B, & a \leqslant s < t \leqslant b, \\ -Ce^{(t-a)A}Pe^{-(s-a)A}B, & a \leqslant t < s \leqslant b. \end{cases}$$

Here $P = e^{aA}\{(N_1 e^{aA} + N_2 e^{bA})^{-1} N_2 e^{bA}\} e^{-aA}$ is the so-called (see [2]) *canonical boundary-value operator* of θ. In what follows we denote the system θ by the quintet $(A, B, C; N_1, N_2)$.

We consider three important classes of systems. The first one consists of *controllable* and *observable* systems. It is known (see [12], [13]) that the notions of controllability and observability do not depend on the boundary conditions. Thus, as in the *causal case* (i.e., when $N_1 = I$, $N_2 = 0$, and hence $P = 0$), a system $\theta = (A, B, C; N_1, N_2)$ is controllable if and only if $\operatorname{Im}[B \; AB \; \cdots \; A^{n-1}B] = X$ (with $n = \dim X$) and θ is observable if and only if $\bigcap_{j=1}^{n} \operatorname{Ker} CA^{j-1} = (0)$.

In order to define the second class we need the operation of *reduction*. Assume that the state space X of $\theta = (A, B, C; N_1, N_2)$ decomposes into a direct sum $X = X_1 + X_0 + X_2$ such that the corresponding partitionings of the operators A, B, C, N_1, and N_2 are as follows:

$$A = \begin{pmatrix} * & * & * \\ 0 & A_0 & * \\ 0 & 0 & * \end{pmatrix}, \qquad B = \begin{pmatrix} * \\ B_0 \\ 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 0 & C_0 & * \end{pmatrix},$$

$$N_\nu = \begin{pmatrix} * & * & * \\ 0 & N_\nu^{(0)} & * \\ 0 & 0 & * \end{pmatrix}, \qquad \nu = 1,2.$$

Here the entries $*$ are unspecified and E is some invertible operator on X. The system $\theta_0 = (A_0, B_0, C_0; N_1^{(0)}, N_2^{(0)})$ is called a *reduction* of θ. Reduction leaves invariant many important characteristics of a system, e.g., the input-

output map. We say that a reduction θ_0 is *proper* if $\dim X_0 < \dim X$. The system θ is called *irreducible* if θ does not admit a proper reduction. Our second class of systems consists of all irreducible ones.

The third class we shall consider is the class of minimal systems. The system θ is called *minimal* if among all systems with the same input-output map as θ the dimension of the state space of θ is as small as possible.

The classical results of systems theory state that for causal systems the three classes introduced above coincide (see, e.g., [10, 15, 16] and the books [1, 9, 11, 14]). Thus for a causal system θ the following three statements are equivalent: (i) θ is minimal; (ii) θ is irreducible; (iii) θ is controllable and observable. For systems with arbitrary well-posed boundary conditions the situation is completely different and each of the three classes is different from the others. We present the following two theorems. (In both theorems the symbol $\vee Z_\alpha$ stands for the linear hull of the spaces Z_α with α running through some index set.)

THEOREM 1. *Let $\theta = (A, B, C; N_1, N_2)$ be a time-invariant boundary-value system. Let X be the state space of θ, and P its canonical boundary-value operator. Then θ is irreducible if and only if*

$$\cap\{\operatorname{Ker} CA^{\alpha_1}P^{\alpha_2}P\cdots A^{\alpha_{2r-1}}P^{\alpha_{2r}} \mid (\alpha_1,\ldots,\alpha_{2r}) \in \Omega(X)\} = (0),$$

$$\vee\{\operatorname{Im} A^{\alpha_1}P^{\alpha_2}\cdots A^{\alpha_{2r-1}}P^{\alpha_{2r}}B \mid (\alpha_1,\ldots,\alpha_{2r}) \in \Omega(X)\} = X.$$

Here $\Omega(X)$ stands for the set of all tuples $(\alpha_1,\ldots,\alpha_{2r})$ of nonnegative integers for which $2r \leqslant \dim X$ and $\sum_{j=1}^{2r}\alpha_j < \dim X$.

THEOREM 2. *Let $\theta = (A, B, C; N_1, N_2)$ be a time-invariant boundary-value system. Let X be the state space of θ, and P its canonical boundary-value operator. Then θ is minimal if and only if*

$$\bigcap_{\nu=0}^{1}\bigcap_{i,j=0}^{n-1} \operatorname{Ker} CA^iP^\nu A^j = (0), \qquad \bigvee_{\nu=0}^{1}\bigvee_{i,j=0}^{n-1} \operatorname{Im} A^iP^\nu A^jB = X,$$

$$\bigcap_{i=0}^{n-1} \operatorname{Ker} CA^i \subset \bigvee_{i=0}^{n-1} \operatorname{Im} A^iB \qquad (n = \dim X).$$

From Theorems 1 and 2 it follows that the class of controllable and observable systems is contained in the class of minimal systems and that the class of

minimal systems is contained in the class of irreducible systems. Simple examples show that in both cases the inclusion is proper.

In the causal case, because minimality is the same as irreducibility, there is a constructive way to reduce a given causal system $\theta = (A, B, C; I, 0)$ to a minimal causal system with the same input-output map as θ. For systems with arbitrary boundary conditions, we give an analogous procedure to reduce a given system to an irreducible one with the same input-output map. However, in general, the resulting system will not be minimal. So, we suggest a much more complicated procedure to get a minimal system without changing the input-output map. To get a controllable and observable system without changing the input-output map is not always possible.

Next we discuss similarity of boundary-value systems. Let $\theta = (A, B, C; N_1, N_2)$ and $\tilde{\theta} = (\tilde{A}, \tilde{B}, \tilde{C}; \tilde{N}_1, \tilde{N}_2)$ be two such systems. Let X and $\tilde{X}$ be the corresponding state spaces. The systems θ and $\tilde{\theta}$ are said to be similar if

$$\tilde{A} = SAS^{-1}, \qquad \tilde{B} = BS^{-1}, \qquad \tilde{C} = SC, \qquad \tilde{N}_i = EN_iS^{-1} \quad (i = 1,2)$$

with some invertible operators $S, E: X \to \tilde{X}$. Clearly, two similar systems have the same input-output map. For minimal causal systems the converse statement holds true. For systems with arbitrary well-posed boundary conditions the situation is more complex, as the following theorems show. In what follows q denotes the maximum of the state-space dimensions of θ and $\tilde{\theta}$.

THEOREM 3. *Assume θ and $\tilde{\theta}$ are both controllable and observable. Then θ and $\tilde{\theta}$ are similar if and only if θ and $\tilde{\theta}$ have the same input-output map, or equivalently,*

$$CA^{2q-1}B = \tilde{C}\tilde{A}^{2q-1}\tilde{B}, \qquad CA^iP^\nu A^jB = \tilde{C}\tilde{A}^i\tilde{P}^\nu\tilde{A}^j\tilde{B}$$

$$(\nu = 0,1; \quad i, j = 0, \ldots, q-1).$$

THEOREM 4. *Assume θ and $\tilde{\theta}$ are both minimal. Then θ and $\tilde{\theta}$ are similar if and only if*

$$CA^iP^\nu A^jP^\alpha A^rP^\mu A^sB = \tilde{C}\tilde{A}^i\tilde{P}^\nu\tilde{A}^j\tilde{P}^\alpha\tilde{A}^r\tilde{P}^\mu\tilde{A}^s\tilde{B},$$

$$CA^iPA^{2q-1}PA^sB = \tilde{C}\tilde{A}^i\tilde{P}\tilde{A}^{2q-1}\tilde{P}\tilde{A}^s\tilde{B}$$

for $\nu, \alpha, \mu = 0, 1$ and $i, j, r, s = 0, \ldots, q$.

THEOREM 5. *Assume θ and $\tilde{\theta}$ are both irreducible. Then θ and $\tilde{\theta}$ are similar if and only if*

$$CA^{\alpha_1}P^{\alpha_2}\cdots A^{\alpha_{2q-1}}P^{\alpha_{2q}}B = \tilde{C}\tilde{A}^{\alpha_1}\tilde{P}^{\alpha_2}\cdots\tilde{A}^{\alpha_{2q-1}}\tilde{P}^{\alpha_{2q}}\tilde{B}$$

for all tuples $(\alpha_1,\ldots,\alpha_{2q})$ of nonnegative integers with $\sum_{j=1}^{2q}\alpha_j \leqslant 2q-1$.

The set of problems discussed in this presentation have been studied before (see [3–6]) in the larger framework of time-varying linear systems with well-posed boundary conditions. When specialized to time-invariant systems, the results of [3–6] differ essentially from those presented in this work, because here minimality and irreducibility are considered in the smaller class of time-invariant systems. Note that the minimality theorem for stationary boundary-value systems in [13] can be easily derived as a corollary of Theorem 2.

The proofs of Theorems 1–5 as well as specifications for stationary and displacement systems and some results on stability of minimal systems can be found in [7]. Minimality and irreducibility of boundary-value systems are also discussed in [8] in a general framework of minimal realizations for matrix functions in several variables.

REFERENCES

1 H. Bart, I. Gohberg, and M. A. Kaashoek, Minimal factorization of matrix and operator functions, in *Operator Theory: Advances and Applications*, Vol. 1, Birkhäuser, Basel, 1979.

2 I. Gohberg and M. A. Kaashoek, Time varying linear systems with boundary conditions and integral operators, I. The transfer operator and its properties, *Integral Equations Operator Theory* 7:325–391 (1984).

3 I. Gohberg and M. A. Kaashoek, Time varying linear systems with boundary conditions and integral operators, II. Similarity and reduction, Report 261, Dept. of Mathematics and Computer Science, Vrije Univ., Amsterdam, 1984.

4 I. Gohberg and M. A. Kaashoek, Similarity and reduction for time varying linear systems with well-posed boundary conditions, *SIAM J. Control Optim.*, to appear.

5 I. Gohberg and M. A. Kaashoek, Minimal representations of semi-separable kernels and systems with separable boundary conditions, to appear.

6 I. Gohberg and M. A. Kaashoek, On minimality and stable minimality of time varying linear systems with well-posed boundary conditions, *Internat. J. Control*, to appear.

7 I. Gohberg, M. A. Kaashoek, and L. Lerer, Minimality and irreducibility of time invariant linear boundary value systems, *Internat. J. Control*, to appear.

8 I. Gohberg, M. A. Kaashoek, and L. Lerer, Minimal realizations for some classes of rational matrix functions in several variables and linear systems, in preparation.

9 T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, N. J., 1980.
10 R. E. Kalman, Mathematical description of linear dynamical systems, *SIAM J. Control* 1(2):152–192 (1963).
11 R. E. Kalman, P. L. Falb, and M. A. Arbib, *Topics in Mathematical Systems Theory*, McGraw-Hill, New York, 1969.
12 A. J. Krener, *Boundary Value Linear Systems*, 75/76, Astérisque, 1980, pp. 149–165.
13 A. J. Krener, Acausal realization theory, Part I: Linear deterministic systems, to appear.
14 H. H. Rosenbrock, *State Space and Multivariable Theory*, Nelson, London, 1970.
15 L. Weiss, On the structure theory of linear differential systems, *SIAM J. Control* 6:659–680 (1968).
16 D. C. Youla, The synthesis of linear dynamical systems from prescribed weighting patterns, *SIAM J. Appl. Math.* 14:527–249 (1966).

MULTIPLICATIVITY AND MIXED MULTIPLICATIVITY FOR OPERATOR NORMS AND MATRIX NORMS

by MOSHE GOLDBERG[20]

Let $\mathbf{V}$ be a normed, finite- or infinite-dimensional vector space over the complex field $\mathbf{C}$, and let $\mathscr{B}(\mathbf{V})$ be the algebra of bounded linear operators on $\mathbf{V}$. A real-valued function

$$N: \mathscr{B}(\mathbf{V}) \to \mathbf{R}$$

is called a *norm* on $\mathscr{B}(\mathbf{V})$ if for all A, $B \in \mathscr{B}(\mathbf{V})$ and $\alpha \in \mathbf{C}$,

$$N(A) > 0, \qquad A \neq 0,$$

$$N(\alpha A) = |\alpha| \cdot N(A),$$

$$N(A + B) \leqslant N(A) + N(B).$$

If in addition N is multiplicative, i.e.,

$$N(AB) \leqslant N(A)N(B) \qquad \forall A, B \in \mathscr{B}(\mathbf{V}),$$

[20] Department of Mathematics, Technion–Israel Institute of Technology, Haifa 32000, Israel. Research sponsored in part by the Air Force Office of Scientific Research, Grant AFOSR-83-0150.

we say that N is an *operator norm* on $\mathscr{B}(\mathbf{V})$. If $\mathscr{B}(\mathbf{V})$ is an algebra of (finite) matrices and N is multiplicative, then N is called a *matrix norm.*

The first multiplicative example that comes to mind is of course, the ordinary operator norm

$$\|A\| = \sup\{|Ax| : x \in \mathbf{V},\ |x| = 1\}, \tag{1}$$

where $|\cdot|$ is the vector norm on $\mathbf{V}$.

If $\mathbf{V}$ is a (finite- or infinite-dimensional) Hilbert space, then perhaps the best known example of a nonmultiplicative norm on $\mathscr{B}(\mathbf{V})$ is the numerical radius (e.g., [9, 2, 8])

$$r(A) = \sup\left\{|(Ax, x)| : x \in \mathbf{V},\ |x| = (x, x)^{1/2} = 1\right\}, \tag{2}$$

where $(\cdot, \cdot)$ is the inner product on $\mathbf{V}$.

Another example of considerable interest is the l_p norm, $1 \leqslant p \leqslant \infty$, of an $n \times n$ matrix $A = (\alpha_{ij}) \in \mathbf{C}_{n\times n}$:

$$|A|_p = \left(\sum_{i,j=1}^{n} |\alpha_{ij}|^p\right)^{1/p}. \tag{3}$$

Ostrowski [11] has shown that this norm is multiplicative (i.e., a matrix norm) if and only if $1 \leqslant p \leqslant 2$.

Given a norm N on $\mathscr{B}(\mathbf{V})$ and a fixed constant $\mu > 0$, then obviously $N_\mu \equiv \mu N$ is a norm too. Clearly, N_μ may or may not be multiplicative. If it is, then we call μ a *multiplicativity factor* for N. That is, μ is a multiplicativity factor for N if and only if

$$N(AB) \leqslant \mu N(A)N(B) \qquad \forall A, B \in \mathscr{B}(\mathbf{V}).$$

Having this definition one can easily prove:

THEOREM 1 [5, Theorem 2.1]. *Let N be a norm on $\mathscr{B}(\mathbf{V})$. Then*

(i) *N has multiplicativity factors if and only if*

$$\mu_{\min} \equiv \sup\{N(AB) : N(A) = N(B) = 1,\ A, B \in \mathscr{B}(\mathbf{V})\} < \infty. \tag{4}$$

(ii) *If $\mu_{\min} < \infty$, then μ is a multiplicativity factor for N if and only if $\mu \geqslant \mu_{\min}$.*

In the finite-dimensional case, compactness immediately implies that $\mu_{\min} < \infty$; hence N always has multiplicativity factors. In the infinite-dimensional case, however, N may fail to have multiplicativity factors, as indicated in Example 2.4 of [5].

While Theorem 1 seems to settle the question of characterizing multiplicativity factors, the quantity $\mu_{\min}$ in (4) is often difficult to compute. A more practical approach towards verifying whether a constant $\mu_{\min} > 0$ is the best (least) multiplicativity factor for a given norm N is implied by the following obvious observation:

A constant $\mu_{\min} > 0$ is the best (least) multiplicativity factor for N if

$$N(AB) \leqslant \mu_{\min} N(A) N(B) \qquad \forall A, B \in \mathscr{B}(\mathbf{V}),$$

with equality for some $A = A_0$, $B = B_0$.

With this observation in mind, it was shown by Holbrook [10, Section 2] (and independently in [4, Theorem 10]) that if $\mathbf{V}$ is a Hilbert space of dimension at least 2, and if r is the numerical radius in (2), then μr is an operator norm on $\mathscr{B}(\mathbf{V})$ if and only if $\mu \geqslant 4$; i.e., the best multiplicativity factor for r is $\mu_{\min} = 4$.

Similarly, Goldberg and Straus [6, Corollary 1.1] have shown that the best multiplicativity factor for the l_p norm on $\mathbf{C}_{n\times n}$ defined in (3) is

$$\mu_{\min} = \begin{cases} 1, & 1 \leqslant p \leqslant 2, \\ n^{1-2/p}, & 2 \leqslant p \leqslant \infty. \end{cases}$$

Often, when $\mu_{\min}$ remains unknown, one may obtain multiplicativity factors via the following somewhat stronger version of a result by Gastinel:

THEOREM 2 [1; 4, Theorem 5]. *Let N and M be a norm and an operator norm on $\mathscr{B}(\mathbf{V})$, respectively; and let $\eta \geqslant \xi > 0$ be constants such that*

$$\xi M(A) \leqslant N(A) \leqslant \eta M(A) \qquad \forall A \in \mathscr{B}(\mathbf{V}).$$

Then any μ with $\mu \geqslant \eta/\xi^2$ is a multiplicativity factor for N.

This result was utilized by Goldberg and Straus [4, 5, 7] to obtain multiplicativity factors for certain generalizations of the numerical radius, called C-numerical radii.

The above concepts of multiplicativity and multiplicativity factors can be extended as follows:

DEFINITION 1. Let $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{W}$ be normed vector spaces over $\mathbf{C}$; and let $\mathscr{B}_1 = \mathscr{B}(\mathbf{U},\mathbf{W})$, $\mathscr{B}_2 = \mathscr{B}(\mathbf{V},\mathbf{W})$, and $\mathscr{B}_3 = \mathscr{B}(\mathbf{U},\mathbf{V})$ be the spaces of bounded linear operators from $\mathbf{U}$ into $\mathbf{W}$, $\mathbf{V}$ into $\mathbf{W}$, and $\mathbf{U}$ into $\mathbf{V}$, respectively. If N_1, N_2, and N_3 are norms on $\mathscr{B}_1$, $\mathscr{B}_2$, and $\mathscr{B}_3$, respectively, and $\mu > 0$ is a constant such that

$$N_1(AB) \leqslant \mu N_2(A) N_3(B) \qquad \forall A \in \mathscr{B}_2,\ B \in \mathscr{B}_3,$$

then we say that μ is a *multiplicativity factor for N_1 with respect to N_2 and N_3*.

In analogy with Theorem 1 we have now:

THEOREM 3 (Compare [3, Theorem 1.2]). *Let N_1, N_2, and N_3 be norms as in Definition* 1. *Then*:

(i) *N_1 has multiplicativity factors with respect to N_2 and N_3 if and only if*

$$\mu_{\min} \equiv \sup\{N_1(AB) : N_2(A) = N_3(B) = 1,\ A \in \mathscr{B}_2,\ B \in \mathscr{B}_3\} < \infty.$$

(ii) *If $\mu_{\min} < \infty$, then μ is a multiplicativity factor for N_1 with respect to N_2 and N_3 if and only if $\mu \geqslant \mu_{\min}$.*

We observe, of course, that *a constant $\mu_{\min} > 0$ is the best (least) multiplicativity factor for N_1 with respect to N_2 and N_3 if*

$$N_1(AB) \leqslant \mu_{\min} N_2(A) N_3(B) \qquad \forall A \in \mathscr{B}_2,\ B \in \mathscr{B}_3,$$

with equality for some $A = A_0$, $B = B_0$.

For example, if $\mathbf{V}$ is a Hilbert space, and if $\|\cdot\|$ and r are the operator norm and numerical radius in (1) and (2), then one can show that

$$r(AB) \leqslant 2r(A)\|B\|, \qquad \forall A, B \in \mathscr{B}(\mathbf{V}),$$

with equality for

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \oplus \mathrm{O}, \qquad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \oplus \mathrm{O}.$$

Thus, $\mu_{\min} = 2$ is the best multiplicativity factor for r with respect to r and $\|\cdot\|$.

This example employs only a single vector space and two norms. In order to demonstrate the idea of mixed multiplicativity to its full extent, consider, for $1 \leqslant p \leqslant \infty$, the l_p norm of an $m \times n$ matrix $A = (\alpha_{ij}) \in \mathbf{C}_{m\times n}$:

$$|A|_p = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |\alpha_{ij}|^p \right)^{1/p}.$$

Then, for any (independent) p, q, r with $1 \leqslant p, q, r \leqslant \infty$, Goldberg [3, Theorem 2.1] has shown that the best multiplicativity factor for the l_p norm on $\mathbf{C}_{m\times n}$ with respect to the l_q norm on $\mathbf{C}_{m\times k}$ and the l_r norm on $\mathbf{C}_{k\times n}$ is

$$\mu_{\min} = \lambda_{pq}(m)\lambda_{pr}(n)\lambda_{q'r}(k),$$

where q' (the conjugate of q) satisfies $1/q + 1/q' = 1$, and

$$\lambda_{pq}(m) = \begin{cases} 1, & p \geqslant q, \\ m^{1/p - 1/q}, & q \geqslant p. \end{cases}$$

Evidently, for $p = 1$, $r = q'$, and $m = n = 1$, this result reduces to Hölder's inequality.

REFERENCES

1 N. Gastinel, *Linear Numerical Analysis*, Academic, New York, 1970.
2 M. Goldberg, Certain finite dimensional numerical ranges and numerical radii, *Linear and Multilinear Algebra* 7:329–342 (1979).
3 M. Goldberg, Mixed multiplicativity and l_p norms for matrices, *Linear Algebra Appl.* 73:123–131 (1986).
4 M. Goldberg and E. G. Straus, Norm properties of C-numerical radii, *Linear Algebra Appl.* 24:113–131 (1979).
5 M. Goldberg and E. G. Straus, Operator norms, multiplicativity factors, and C-numerical radii, *Linear Algebra Appl.* 43:137–159 (1982).
6 M. Goldberg and E. G. Straus, Multiplicativity of l_p norms for matrices, *Linear Algebra Appl.* 52/53:351–360 (1983).
7 M. Goldberg and E. G. Straus, Multiplicativity factors for C-numerical radii, *Linear Algebra Appl.* 54:1–16 (1983).
8 M. Goldberg and E. Tadmor, On the numerical radius and its applications, *Linear Algebra Appl.* 42:263–284 (1982).
9 P. R. Halmos, *A Hilbert Space Problem Book*, Van Nostrand, New York, 1967.
10 J. A. R. Holbrook, Multiplicativity properties of the numerical radius in operator theory, *J. Reine Angew. Math.* 237:166–174 (1969).
11 A. Ostrowski, Über Normen von Matrizen, *Math. Z.* 63:2–18 (1955).

MATRICES WITH SIGN-SYMMETRIC DIAGONAL SHIFTS OR SCALAR SHIFTS

by DANIEL HERSHKOWITZ,[21,22]*, VOLKER MEHRMANN,[23] AND HANS SCHNEIDER[21,22]

A square complex matrix is said to be sign-symmetric [weakly sign-symmetric] if it has nonnegative products of symmetrically located minors [almost principal minors].

Weakly sign-symmetric matrices were studied first by Gantmacher and Krein [5, p. 111] and by Kotelianskii [10]. That is why these matrices are also called GKK matrices, e.g. by Fan [4]. One reason for the interest in these classes of matrices is that they contain the important classes of the Hermitian matrices, the totally nonnegative matrices, and the M-matrices. Another reason is the strong linkage between weak sign symmetry and the Fischer-Hadamard determinantal inequalities. This connection is studied in Gantmacher and Krein [5], Koteljanskii [9], Carlson [1], Green [6], and Hershkowitz and Berman [7].

A sufficient condition for positivity of the principal minors of a weakly sign-symmetric matrix in terms of leading principal minors is given by Koteljanskii [10].

Relations between weakly sign-symmetric matrices and ω-matrices are discussed in Engel and Schneider [3] and in Hershkowitz and Berman [8].

Sign symmetry and weak sign symmetry are also related to stability. It was proved by Carlson [2] that sign-symmetric matrices whose principal minors are positive are stable, i.e., their spectra lie in the open right half plane. The same result is conjectured to hold for weakly sign-symmetric matrices too.

In this paper we generalize the concepts of sign symmetry and weak sign symmetry. We define k-sign-symmetric matrices, where k is a nonnegative integer. In view of our definition an $n \times n$ sign-symmetric matrix is a k-sign-symmetric matrix whenever $k \geq (n-1)/2$. The 1-sign-symmetric matrices are those weakly sign-symmetric matrices whose principal minors are real. Since reality of principal minors is assumed in all the results on weakly sign-symmetric matrices quoted above, one may as well consider those as assertions on 1-sign-symmetric matrices.

After giving graph-theoretic preliminaries, we characterize the matrices all of whose diagonal shifts are k-sign-symmetric, that is, matrices A such that $A + D$ is k-sign-symmetric for every real diagonal matrix D. Given a positive k, we show that an irreducible matrix satisfies this condition if and only if it is

[21] Department of Mathematics, Technion, I.I.T., Haifa 32000, Israel.

[22] Department of Mathematics, University of Wisconsin, Madison, Madison, WI 53706.

[23] Fakultät für Mathematik, Universität Bielefeld, Postfach 8640, D-4800 Bielefeld 1, F.R.G.

diagonally similar to a Hermitian matrix. Thus, a matrix satisfies the above shift condition for some positive k if and only if it satisfies the condition for every positive k.

For $k \geqslant 2$, we prove a similar result for a matrix A all of whose scalar shifts $A + tI$, where t is real, are k-sign-symmetric. If $k = 1$, then we need an additional graph-theoretic hypothesis, namely, the reversibility of the chordless directed circuits of even length in the directed graph of A.

The extensions of our results to reducible matrices follow from a theorem that a matrix A is k-sign-symmetric if and only if every diagonal block in the Frobenius normal form of A is k-sign-symmetric.

REFERENCES

1 D. Carlson, Weakly sign-symmetric matrices and some determinantal inequalities, *Colloq. Math.* 27:123–129 (1967).
2 D. Carlson, A class of positive stable matrices, *J. Res. Natl. Bur. Standards* 78B:1–2 (1974).
3 G. M. Engel and H. Schneider, The Hadamard-Fischer inequality for a class of matrices defined by eigenvalue monotonicity, *Linear and Multilinear Algebra* 4:155–176 (1976).
4 K. Fan, Subadditive functions on a distributive lattice and an extension of Szasz's inequality, *J. Math. Anal. Appl.* 18:262–268 (1967).
5 F. R. Gantmacher and M. G. Krein, *Oszillationsmatrizen, Oszillationskerne und kleine Schwingungen mechanischer Systeme*, Akademie Verlag, Berlin, 1960.
6 B. M. Green, On weakly sign-symmetric matrices, *Linear Algebra Appl.* 8:355–360 (1974).
7 D. Hershkowitz and A. Berman, Necessary conditions and a sufficient condition for the Fischer-Hadamard inequalities, *Linear Multilinear Algebra* 13:67–72 (1983).
8 D. Hershkowitz and A. Berman, Notes on ω- and τ-matrices, *Linear Algebra Appl.* 58:169–183 (1984).
9 D. M. Koteljanskii, The theory of nonnegative and oscillating matrices, *Amer. Math. Soc. Transl. Ser.* 2 27:1–8 (1963).
10 D. M. Koteljanskii, A property of sign symmetric matrices, *Amer. Math. Soc. Transl. Ser.* 2 27:19–24 (1963).

EQUALITY CLASSES OF MATRICES

by DANIEL HERSHKOWITZ[24] and HANS SCHNEIDER[24]*

Let A be a complex $n \times n$ matrix, and define the absolute-value matrix $B = |A|$ of A by $b_{ij} = |a_{ij}|$, $i, j = 1, \ldots, n$. Let $\rho(A)$ be the spectral radius of A.

[24] Mathematics Department, University of Wisconsin—Madison, Madison, WI 53706, and Mathematics Department, Technion–Israel Institute of Technology, Haifa 32000, Israel.

Let $\mathscr{U}$ be the set of all noncomplex matrices A such that $\rho(|A|) < 1$. In [6] Ostrowski proved the now very well-known result that, for $A \in \mathscr{U}$,

$$\left|(I - A)^{-1}\right| \leqslant (1 - |A|)^{-1}, \tag{1.1}$$

where the inequality is entrywise.

In [5] Neumaier showed that for $A \in \mathscr{J}$, the set of $n \times n$ irreducible matrices $A \in \mathscr{U}$,

$$\left|(I - A)^{-1}\right| = (I - |A|)^{-1} \tag{1.2}$$

if and only if

$$\text{All circuit products of } A \text{ are positive.} \tag{1.3}$$

It is well known [1,2] that for irreducible A, (1.3) is equivalent to

$$\begin{aligned} &A \text{ is diagonally similar to } |A|, \text{ i.e., there exists a diagonal matrix} \\ &X \text{ such that } A = X|A|X^{-1}. \end{aligned} \tag{1.4}$$

Neumaier shows that the condition

$$\left|(I - A^{-1})\right|_{ij} = \left(I - |A|^{-1}\right)_{ij} \qquad \text{for some } i, j \quad 1 \leqslant i, j \leqslant n, \tag{1.5}$$

which is apparently weaker than (1.2), is in fact equivalent to (1.2)–(1.4) for $A \in \mathscr{J}$. (We have stated special cases of the results of Ostrowski and Neumaier, from which however the general theorems may easily be derived.)

In this paper we generalize Neumaier's results in various directions. We consider the equality (1.2) for general $A \in \mathscr{U}$, omitting the requirement of irreducibility. We use the concept of a 2-twisted chain of the graph $G(A)$ of A, which was defined in [4]. Intuitively, a chain in a directed graph is obtained by putting a pointer at a vertex and moving it in the direction or against the direction of a connected sequence of arcs to another vertex. Each change in direction is a twist. A 2-twisted chain (e.g., cycle) is a chain with at most two twists. Thus a circuit (directed cycle) is a special case of a 2-twisted cycle. We show that for $A \in \mathscr{U}$ the condition (1.2) is equivalent to

$$\begin{aligned} &\text{All cycle products of } A \text{ corresponding to 2-twisted cycles are} \\ &\text{positive.} \end{aligned} \tag{1.6}$$

(and other conditions). This generalizes (1.3).

If C is an $s \times s$ matrix and A is an $n \times n$ matrix, where $s \leqslant n$, we generalized both the Kronecker and Hadamard products in [3] by defining the $n \times n$ matrix $C \times\times A$. Thus, if A is partitioned into s^2 matrices A_{ij}, $i, j = 1, \ldots, s$, then $C \times\times A$ is the matrix whose blocks are $c_{ij}A_{ij}$, $i, j = 1, \ldots, s$. Here we show that if $A \in \mathscr{U}$ is in Frobenius normal form, then A satisfies (1.2) if and only if

> A is diagonally similar to $C \times\times |A|$, where C is an upper triangular $s \times s$ matrix ($s \leqslant n$) such that $|c_{ij}|$ is 1 or 0, and c_{ii} is 1 or 0, $i, j = 1, \ldots, s$, and zC satisfies (1.2) for $0 < z < 1$. (1.7)

This generalizes (1.4).

We also generalize (1.5) by defining the concept of a $G(A)$-access cover. A subset Γ of $\langle n \rangle \times \langle n \rangle$, where $\langle n \rangle = \{1, \ldots, n\}$, is a $G(A)$-access cover if for each $(i, j) \in \langle n \rangle \times \langle n \rangle$ there is an $(h, k) \in \Gamma$ such that h has access to i in $G(A)$ and j has access to k in $G(A)$. We observe that $\{(i, j)\}$ is a $G(A)$-access cover for all $(i, j) \in \langle n \rangle \times \langle n \rangle$ if and only if A is irreducible [or equivalently, $G(A)$ is strongly connected]. Thus, if Γ is a $G(A)$-access cover and $A \in \mathscr{U}$, then (1.2) is equivalent to

$$\left|(I - A)^{-1}\right|_{ij} = (I - |A|)^{-1}_{ij} \qquad \text{for} \quad (i, j) \in \Gamma.$$

It is easily seen that (1.2) is equivalent to

$$\left| \sum_{s \in N} A^s \right| = \sum_{s \in N} |A|^s \tag{1.8}$$

for $A \in \mathscr{U}$, where N is the set of natural numbers. Since, for all subsets S of N,

$$\left| \sum_{s \in S} A^s \right| \leqslant \sum_{s \in S} |A|^s, \tag{1.9}$$

it is natural to define $\mathrm{Equ}(\mathscr{A}, \Gamma, S)$ to be the set of all $A \in \mathscr{A}$ such that

$$\left| \sum_{s \in S} A^s \right| = \sum_{s \in S} |A|^s \qquad \text{for} \quad (i, j) \in \Gamma, \tag{1.10}$$

where $\mathscr{A} \subseteq \mathscr{U}$, $\Gamma \subseteq \langle n \rangle \times \langle n \rangle$, and $S \subseteq N$.

The equivalences stated above, and others, are stated in terms of $\mathrm{Equ}(\mathscr{U}, \Gamma, N)$. It is clear that $\mathrm{Equ}(\mathscr{A}, \Gamma, S) \supseteq \mathrm{Equ}(\mathscr{A}, \Gamma, N)$ for $S \subseteq N$. We therefore call a subset of S of N $(\mathscr{A}, \Gamma)$-sufficient if

$$\mathrm{Equ}(\mathscr{A}, \Gamma, S) = \mathrm{Equ}(\mathscr{A}, \Gamma, N).$$

We give conditions equivalent to $(\mathscr{J}, \langle n \rangle \times \langle n \rangle)$-sufficiency and $(\mathscr{U}, \langle n \rangle \times \langle n \rangle)$-sufficiency. One of our results shows that if S is a subset of N, then S is $(\mathscr{J}, \langle n \rangle \times \langle n \rangle)$-sufficient if and only if $\langle n \rangle \subseteq \mathrm{CD}(S)$, where $\mathrm{CD}(S)$ is defined thus: Let $D(S)$ consist of all differences $s - t$, where $s > t$ and $s, t \in S$, and let $\mathrm{CD}(S)$ be the set of all greatest common divisors of subsets of $D(S)$. As an application of this theorem let $n = 10$. If $S = \{3, 9, 10, 13, 18\}$ then $\mathrm{CD}(S) = \langle 10 \rangle \cup \{15\}$ and hence S is $(\mathscr{J}, \langle 10 \rangle \times \langle 10 \rangle)$-sufficient. In other words, if A is an irreducible 10×10 matrix, then the equality

$$|A^3 + A^9 + A^{10} + A^{13} + A^{18}| = |A|^3 + |A|^9 + |A|^{10} + |A|^{13} + |A|^{18}$$

implies that for every set S of positive integers and arbitrary positive coefficients α_s, $s \in S$, we have

$$\left| \sum_{s \in S} \alpha_s A^s \right| = \sum_{s \in S} \alpha_s |A|^s$$

(assuming that the above series converge or that S is finite). In particular if $\rho(|A|) < 1$ then

$$\left| (I - A)^{-1} \right| = (I - |A|)^{-1}$$

The general problem of characterizing $(\mathscr{A}, \Gamma)$-sufficient sets and minimal $(\mathscr{A}, \Gamma)$-sufficient sets for $\mathscr{A} \subseteq \mathscr{U}$ and $\Gamma \subseteq \langle n \rangle \times \langle n \rangle$ is open.

REFERENCES

1 G. M. Engel and H. Schneider, Cyclic and diagonal products on a matrix, *Linear Algebra Appl.* 7:301–335 (1973).

2 M. Fiedler and V. Pták, Cyclic products and an inequality for determinants, *Czechoslovak Math. J.* 19:428–450 (1969).

3 S. Friedland, D. Hershkowitz, and H. Schneider, Matrices whose powers are *M*-matrices or *Z*-matrices, to appear.

4 D. Hershkowitz and H. Schneider, On $2k$-twisted graphs, *European J. Combin.*, to appear.

5 A Neumaier, The extremal case of some matrix inequalities, *Arch. Math. (Basel)* 43:137–141 (1984).
6 A. Ostrowski, Über die Determinanten mit überwiegender Hauptdiagonale, *Comment. Math. Helv.* 10:69–96 (1937).

PERTURBED AND MIXED TOEPLITZ MATRICES AS GENERALIZATION OF THE RESULTANT MATRIX

by B. A. KON[25]

Let λ_j $(j = 1,2,\ldots,\nu)$ denote the distinct common zeros of the polynomials $a(z) = \sum_{i=0}^{n} a_i z^i$ and $b(z) = \sum_{i=0}^{m} b_i z^i$ with complex coefficients, and let ρ_j be the common multiplicity of λ_j. Denote $N(a,b) = \sum_{j=1}^{\nu} \rho_j$.

It is well known (see [1] and references therein) that the vectors

$$\varphi_{i,k} = \left\{ \binom{p}{k} \lambda_i^{p-k} \right\}_{p=0}^{m+n-1} \qquad (k = 0,1,\ldots,\rho_i - 1, \quad i = 1,2,\ldots,\nu)$$

form a basis of the kernel of the resultant (Sylvester) matrix

$$R(a,b) = \begin{pmatrix} a_0 & \cdots & a_n & & \\ & \ddots & & \ddots & \\ & & a_0 & \cdots & a_n \\ b_0 & \cdots & b_m & & \\ & \ddots & & \ddots & \\ & & b_0 & \cdots & b_m \end{pmatrix}.$$

The main aim of this note is to extend this result to some classes of analytic functions. As pointed out by Gohberg and Lerer ([2],[3]; see also [4],[5]), a generalization of the resultant matrix in the case of analytic functions leads to a convolution-type operator. The properties of this resultant are shown to be strongly dependent on the form of the domain.

I. Functions Holomorphic in a Unit Disk

Let $a(z)$ and $b(z)$ be analytic functions in the unit disk $D_v^+ = \{z \in \mathbb{C} : |z| < 1\}$, and let them admit the following representations as absolutely

[25] Department of Mathematics, Technion–Israel Institute of Technology, Haifa, Israel.

convergent series on D_v^+: $a(z)=\sum_{i=0}^{\infty}a_i z^i$, $b(z)=\sum_{i=0}^{\infty}b_i z^i$. Let $N(b)$ be the number of all zeros of $b(z)$ in D_v^+, and let λ_j $(j=1,2,\ldots,\nu)$ denote all the distinct common zeros of $a(z)$ and $b(z)$ in D_v^+ with the common multiplicity ρ_j, and $N(a,b)=\sum_{j=1}^{\nu}\rho_j$. Introduce in the space l^2 the operator $K_n(a,b)$ by the matrix

$$K_n(a,b)=\begin{pmatrix} a_0 & a_1 & a_2 & \cdots & & \\ & \ddots & \ddots & \ddots & & \\ & & a_0 & a_1 & a_2 & \cdots \\ b_0 & b_1 & b_2 & \cdots & & \\ & b_0 & b_1 & b_2 & \cdots & \\ & & \ddots & \ddots & \ddots & \end{pmatrix} \Bigg\} n \text{ rows}$$

THEOREM 1. *Assume* $b(z)\neq 0$, $|z|=1$. *For every* $n\geq N(b)$ *the vectors*

$$\varphi_{i,j}=\left\{\binom{i+k}{k}\lambda_j^{-(i+k+1)}\right\}_{k=0}^{\infty} \qquad (j=1,2,\ldots,\nu,\quad i=0,1,\ldots,\rho_j-1) \quad (1)$$

constitute a basis of $\operatorname{Ker}K_n(a,b)$. *In particular,* $\dim\operatorname{Ker}K_n(a,b)=N(a,b)$.

Note that this theorem in a more general setting is established in [5].

We also remark that some difficulty in constructing the operator $K_n(a,b)$ in Theorem 1 is caused by the need to know the number $N(b)$ of distinct zeros of $b(z)$. This is removed in the following proposition.

Construct the operator $R(a,b)$ in l^2 by the following "mixed" matrix:

$$R(a,b)=\begin{bmatrix} a_0 & a_1 & a_2 & a_3 & \cdots \\ b_0 & b_1 & b_2 & b_3 & \cdots \\ & a_0 & a_1 & a_2 & \cdots \\ & b_0 & b_1 & b_2 & \cdots \\ & & a_0 & a_1 & \cdots \\ & & b_0 & b_1 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

THEOREM 2 [6]. *Assume* $a(z)\neq 0$, $b(z)\neq 0$ *on* $|z|=1$. *Then* $\dim\operatorname{Ker}R(a,b)=N(a,b)$, *and the vectors in* (1) *form a basis of the nullspace of* $R(a,b)$.

II. Functions Analytic in an Annulus

We now assume that $a(z)$ and $b(z)$ are analytic functions in the annulus $D^+ = \{z \in \mathbb{C} : R^{-1} < |z| < R,\ R > 1\}$ and admit the following representation in the absolutely convergent series on D^+: $a(z) = \sum_{j=-\infty}^{\infty} a_j z^j$, $b(z) = \sum_{j=-\infty}^{\infty} b_j z^j$. Denote

$$l^2(R) = \left\{ \varphi = \{\varphi_j\}_{j=-\infty}^{\infty} \in l^2 : \sum_{j=-\infty}^{\infty} |\varphi_j| R^{-|j|} < \infty \right\},$$

and introduce on $l^2(R)$ the operator $R_{m,\alpha}(a,b)$ by the matrix

$R_{m,\alpha}(a,b)$

$$= \left(\begin{array}{ccccccccccc}
\cdots & a_2 & a_1 & a_0 & a_{-1} & a_{-2} & \cdots & & & & \\
 & \cdots & a_2 & a_1 & a_0 & a_{-1} & a_{-2} & \cdots & & & \\
 & & \cdots & b_2 & b_1 & b_0 & b_{-1} & b_{-2} & \cdots & & \\
 & & & \ddots & \ddots & \ddots & \ddots & \ddots & & & \\
 & & & & \cdots & b_2 & b_1 & b_0 & b_{-1} & b_{-2} & \cdots \\
 & & & & & \cdots & a_{\alpha+2} & a_{\alpha+1} & a_\alpha & a_{\alpha-1} & \cdots \\
 & & & & & & \cdots & a_{\alpha+2} & a_{\alpha+1} & a_\alpha & \cdots
\end{array} \right) \Bigg\} m \text{ rows}$$

THEOREM 3 [6]. *Let $a(z)$ and $b(z)$ be as before, and $a(z) \neq 0$ for $z \in \partial D^+$. Then for all $\alpha > m$ and $m \geqslant N(a) - N(a,b)$ the dimension of the subspace* $\operatorname{Ker} R_{m,\alpha}(a,b)$ *is equal to the number of common zeros (counting multiplicities of $a(z)$ and $b(z)$) in D^+.*

Recalling the remark following Theorem 1, define a "mixed" resultant operator in $l^2(R)$ for n functions $a_i(z) = \sum_{k=-\infty}^{\infty} a_{i,k} z^k$ $(i = 1,2,\ldots,n)$ that are analytic in D^+. Denote $A_k = [a_{1,k}, a_{2,k}, \ldots, a_{n,k}]^T$ $(k = 0, \pm 1, \pm 2, \ldots)$. The relation

$$R(a_1, a_2, \ldots, a_n) = \|A_{j-k}\|_{j,k=-\infty}^{\infty}$$

defines an operator on $l^2(R)$.

THEOREM 4 [6]. *Assume $a_i(z) \neq 0$, $z \in \partial D^+$ $(i = 1,2,\ldots,n)$. Then the vectors*

$$\varphi_{i,j} = \left\{ \binom{i+k}{k} \lambda_j^{-(i+k+1)} \right\}_{k=-\infty}^{\infty} \qquad (j = 1,2,\ldots,\nu, \quad i = 0,1,\ldots,\rho_j - 1),$$

where λ_j $(j = 1,2,\ldots,\nu)$ *are all common zeros of* $a_1(z), a_2(z),\ldots,a_n(z)$ *in* D^+ *with common multiplicity* ρ_j, *form a basis of* $\operatorname{Ker} R(a_1, a_2,\ldots,a_n)$.

REFERENCES

1 I. Gohberg and G. Heinig, Resultant matrix and its generalization. I. Resultant operator of matrix polynomials, *Acta Sci. Math. (Szeged)* 37:41–61 (1975).
2 I. C. Gohberg and L. E. Lerer, Singular integral operators as a generalization of the resultant matrix, *Applicable Anal.* 7:191–205 (1978).
3 L. C. Gohberg and L. E. Lerer, Resultant operators of a pair of analytic functions, *Proc. Amer. Math. Soc.* 72, No. 1 (1978).
4 B. A. Kon and L. E. Lerer, Resultant operators for analytic functions in a simple connected domain, *Integral Equations Operator Theory*, to appear.
5 B. A. Kon, Resultant operators for analytic functions in a multiply connected domain, *J. Math. Anal. Appl.*, to appear.
6 B. A. Kon, The mixed Toeplitz matrix as resultant operator for analytic functions, to appear.

BEZOUTIAN FOR SEVERAL MATRIX POLYNOMIALS AND POLYNOMIAL LYAPUNOV-TYPE EQUATIONS

by L. LERER[26] and M. TISMENETSKY[27]*

In this work we introduce a new concept of a Bezoutian associated with a family of matrix polynomials and, as an application, extend the classical Lyapunov method for determining the inertia of a matrix to matrix polynomials.

Let $L_r(\lambda)$ and $M_r(\lambda)$ $(r = 1,2,\ldots,s)$ be regular [i.e. $L(\lambda) \not\equiv 0$] $n \times n$ matrix polynomials (we follow the book [5] in presenting the theory of matrix polynomials) such that $\sum_{r=1}^{s} M_r(\lambda)L_r(\lambda) = 0$. The *generalized Bezoutian* is defined as an $\mu n \times \nu n$ block matrix

$$\mathbb{B} = \mathbb{B}(M_1,\ldots,M_s;\ L_1,\ldots,L_s) := \left[\Gamma_{ij}\right]_{i,j=0}^{\mu-1,\nu-1}, \tag{1}$$

[26]Department of Mathematics, Technion–Israel Institute of Technology, Haifa 32000, Israel.
[27]IBM Scientific Center, Technion City, Haifa 32000, Israel.

the entries Γ_{ij} of which are found from the expansion

$$(\lambda - \mu)^{-1} \sum_{r=1}^{s} M_r(\lambda) L_r(\mu) = \sum_{i,j=0}^{\mu-1,\nu-1} \lambda^i \mu^j \Gamma_{ij}.$$

Here $\mu = \max_{1 \leqslant r \leqslant s} \deg M_r(\lambda)$, $\nu = \max_{1 \leqslant r \leqslant s} \deg L_r(\lambda)$. In the scalar case $n = 1$ the classical Bezout matrix of two polynomials $p(\lambda)$ and $q(\lambda)$ is obtained by setting $s = 2$, $M_1(\lambda) = L_2(\lambda) = p(\lambda)$, and $L_1(\lambda) = -M_2(\lambda) = q(\lambda)$. In the matrix case with $s = 2$ the matrix in (1) coincides with the Bezoutian introduced in [1] if $M_2(\lambda)$ is replaced by $-M_2(\lambda)$.

It turns out that the concept of a generalized Bezoutian plays a fundamental role in relating matrix equations and equations in matrix polynomials.

Consider the following equation in $n \times n$ matrix polynomials:

$$M(\lambda)Y(\lambda) + Z(\lambda)L(\lambda) = R(\lambda), \tag{E_1}$$

where $M(\lambda) = \lambda^\mu I + \sum_{j=0}^{\mu-1} \lambda^j m_j$ and $L(\lambda) = \lambda^\nu I + \sum_{j=0}^{\nu-1} \lambda^j l_j$ are monic polynomials and $\deg R \leqslant \nu + \mu - 1$. The set of all solutions pairs $(Y(\lambda), Z(\lambda))$ such that $\deg Y \leqslant \gamma$, $\deg Z \leqslant \delta$ will be denoted by $\mathscr{Y}_{\gamma,\delta}(E_1)$. Without loss of generality we may assume that $R(\lambda)$ is represented in the form

$$R(\lambda) = -\sum_{r=3}^{s} M_r(\lambda) L_r(\lambda), \tag{2}$$

where $\deg M_r \leqslant \mu$, $\deg L_r \leqslant \nu - 1$ $(r = 3, \ldots, s)$.

The representation (2) allows us to associate to each solution $(Y(\lambda), Z(\lambda)) \in \mathscr{Y}_{\nu-1,\mu-1}(E_1)$ the $\mu n \times \nu n$ Bezout matrix $\mathbb{B}_{Y,Z} := \mathbb{B}(M, Z, M_3, \ldots, M_s; Y, L, L_3, \ldots, L_s)$ generated by the equation (E_1).

Now consider the matrix equation

$$\hat{C}_M X - X C_L = R, \tag{E_2}$$

where C_L and $\hat{C}_M$ stand for the first and the second companion matrices of $L(\lambda)$ and $M(\lambda)$, respectively, and

$$R = \sum_{r=3}^{s} M_r(\hat{C}_M, Z^{(0)}) L_r(X^{(0)}, C_L). \tag{3}$$

Here $X^{(0)} = [I\ 0\ \cdots\ 0]$, $Z^{(0)} = [I\ 0\ \cdots\ 0]^T$, and for $A(\lambda) = \sum_{j=0}^{\alpha} \lambda^j a_j$ we set

$$A\big(X^{(0)}, C_L\big) := \sum_{j=0}^{\alpha} a_j X^{(0)} C_L^j, \qquad A\big(\hat{C}_M, Z^{(0)}\big) := \sum_{j=0}^{\alpha} \hat{C}_M^j Z^{(0)} a_j.$$

By $\mathscr{S}(\mathrm{E}_2)$ we denote the set of all solutions of the equation (E_2).

THEOREM 1. *Preserving the above assumptions, the equation* (E_1) *is solvable of and only if the equation* (E_2) *is solvable. Moreover, the mapping* β *defined on* $\mathscr{Y}_{\nu-1,\mu-1}(E_1)$ *by* $\beta(Y(\lambda), Z(\lambda)) = \mathbb{B}_{Y,Z}$ *is injective and maps* $\mathscr{Y}_{\nu-1,\mu-1}(\mathrm{E}_1)$ *onto* $\mathscr{S}(\mathrm{E}_2)$. *The inverse mapping* $\beta^{-1}: \mathscr{S}(\mathrm{E}_2) \to \mathscr{Y}_{\nu-1,\mu-1}(\mathrm{E}_1)$ *acts as follows. If* $S = [S_{jk}]_{j,k=0}^{\mu-1,\nu-1} \in \mathscr{S}(\mathrm{E}_2)$, *then* $\beta^{-1}(S) = (Y_S(\lambda), Z_S(\lambda))$, *where*

$$\begin{aligned} Y_S(\lambda) &= \sum_{k=0}^{\nu-1} \lambda^k S_{\mu-1,k} - \sum_{r=3}^{s} m_{r\mu} L_r(\lambda), \\ Z_S(\lambda) &= -\sum_{j=0}^{\mu-1} \lambda^j S_{j,\nu-1}. \end{aligned} \tag{4}$$

Note that the equation (E_2) depends on the choice of the decomposition of $R(\lambda)$ in (E_1) and that the one-to-one correspondence in the theorem relates to that choice.

Using theorem 1, we describe the set $\mathscr{Y}_{\nu,\mu}(\mathrm{E}_1)$ also in the case $\deg R(\lambda) \leqslant \nu + \mu$ and obtain these results in a more general setting of arbitrary matrix polynomials $L(\lambda)$ and $M(\lambda)$. Also we show that (E_2) can be replaced by matrix equations of more general type of which the coefficients are determined by the spectral data of $L(\lambda)$ and $M(\lambda)$.

Returning to monic polynomials $L(\lambda)$ and $M(\lambda)$, we remark that S. Barnett [2] was probably the first who observed in the case $\deg R(\lambda) \leqslant \nu + \mu - 2$ the connection between the equations (E_1) and (E_2), and, by a different method, established the formulas (4) for this case, i.e. with $m_{r\mu} = 0$, $r = 3, \ldots, s$. However, the fact that the connection between (E_1) and (E_2) is established by the mapping β, as well as its properties, is new even in this case.

Proceeding to applications, recall that the *inertia* of a matrix $A \in \mathbb{C}^{n\times n}$ is defined as a triple of integers $\mathrm{In}\, A = (\pi(A), \nu(A), \delta(A))$, where $\pi(A)$, $\nu(A)$, and $\delta(A)$ denote the numbers of eigenvalues counted with their multiplicities with positive, negative, and zero real parts, respectively. If $\mathrm{In}\, A = (0, n, 0)$,

then the matrix A is called *stable*. Lyaponov's theorem [8] (see also [6], for instance) states that an $n \times n$ complex matrix A is stable if and only if there exists a negative definite matrix H such that for some positive definite W (for brevity, $W > 0$),

$$AH + HA^* = W, \qquad W > 0. \tag{5}$$

The Lyaponov theorem was generalized by M. G. Krein (see [4]) and, independently, by A. Ostrowski and H. Schneider [9], and O. Taussky [10], who show that if a Hermitian matrix H satisfies (5), then

$$\pi(A) = \pi(H), \qquad \nu(A) = \nu(H), \qquad \delta(A) = \delta(H) = 0. \tag{6}$$

Conversely, if $\delta(A) = 0$, then there exists a Hermitian matrix H such that (5) and (6) hold.

More general inertia theorems expressing the inertia of A in terms of that of H, $H = H^*$, satisfying the Lyaponov equation $AH + HA^* = W$ with a nonnegative definite W, are established in [3], [11], and [12].

In this work the classical inertia results are extended to matrix polynomials.

Define the *inertia* of a regular matrix polynomial $L(\lambda)$ by the triple $\gamma(L) = (\gamma_+(L), \gamma_-(L), \gamma_0(L)$, where $\gamma_\pm(L), \gamma_0(L)$ denote the numbers of eigenvalues of $L(\lambda)$, counted with their multiplicities, lying in the open right (left) half plane and on the imaginary axis (including infinity), respectively. It turns out that for matrix polynomials the equation in (5) should be replaced by

$$L(-\lambda)L_1^*(\lambda) + L_1(-\lambda)L^*(\lambda) = \sum_{r=1}^{s} M_r(-\lambda)M_r^*(\lambda), \tag{7}$$

where $\deg L(\lambda) = l$, $\deg L_1(\lambda) \leqslant l-1$, and $\deg M_r(\lambda) \leqslant l-1$ $(1 \leqslant r \leqslant s)$. Here $A^*(\lambda) = \sum_{j=0}^{k} \lambda^j a_j^*$ for $A(\lambda) = \sum_{j=0}^{k} \lambda^j a_j$. Observe that (7) becomes (5) when $L(\lambda) = A - \lambda I$, $L_1(\lambda) \equiv H$, and $M_1 = W^{\frac{1}{2}}$.

When concerned with determining the inertia with respect to the real axis, the equation

$$L^*(\lambda)L_1(\lambda) + L_1^*(\lambda)L_1(\lambda) = \sum_{r=1}^{s} M_r^*(\lambda)M_r(\lambda)$$

should be taken instead of (7).

Denote by $\hat{\mathbb{B}}$ the matrix $\mathbb{B}D$, where $D = \operatorname{diag}[I, -I, \ldots, (-1)^{l-1}I]$ and $\mathbb{B}$ denotes the Bezoutian associated with (7). Note that $\hat{\mathbb{B}}$ is Hermitian. Let $M_r(\lambda) = \sum_{j=0}^{l-1} \lambda^j m_{jr}$ and

$$\sum_{r=1}^{l} \tilde{M}_r^* \tilde{M}_r > 0, \tag{8}$$

where $\tilde{M}_r = [m_{0r}\ im_{1r}\ \cdots\ i^{l-1}m_{l-1,r}]$.

Using Theorem 1, we establish the following generalization of the Krein-Ostrowski-Schneider-Taussky theorem.

THEOREM 2. *Let $L(\lambda)$ be an $n \times n$ monic matrix polynomial. If there exists a matrix polynomial $L_1(\lambda)$ of degree $\leqslant l-1$ such that* (7) *holds for some $M_r(\lambda)$ satisfying* (8), *then*

$$\gamma_+(L) = \pi(\hat{\mathbb{B}}), \qquad \gamma_-(L) = \nu(\hat{\mathbb{B}}), \qquad \gamma_0(L) = \delta(\hat{\mathbb{B}}) = 0. \tag{9}$$

Conversely, if $\gamma_0(L) = 0$, then there is a matrix polynomial $L_1(\lambda)$ of degree $\leqslant l-1$ such that (7)–(9) are valid.

Note that the assertion of Theorem 2 can be extended to an arbitrary regular matrix polynomial. Also, generalizations of other inertia theorems are obtained. The condition (8) is removed in this case.

REFERENCES

1 B. D. O. Anderson and E. I. Jury, Generalized Bezoutian and Sylvester matrices in multivariable linear control, *IEEE Trans. Automat. Control* AC-21:551–556 (1976).

2 S. Barnett, Regular polynomial matrices having relatively prime determinants, *Proc. Cambridge Philos. Soc.* 65:585–590 (1969).

3 D. Carlson and H. Schneider, Inertia theorems for matrices: The semidefinite case, *J. Math. Anal. Appl.* 6:430–446 (1963).

4 J. Daleskii and M. G. Krein, Stability of solutions of differential equations in Banach space, Amer. Math. Soc. Transl. No. 43, 1974.

5 I. Golberg, P. Lancaster, and L. Rodman, *Matrix Polynomials*, Academic, 1982.

6 P. Lancaster and M. Tismenetsky, *Theory of Matrices*, Academic, 1985.

7 L. Lerer and M. Tismenetsky, The eigenvalue-separation problem for matrix polynomials, *Integral Equations and Operator Theory* 5:386–445 (1982).

8 A. M. Lyaponov, Problème général de la stabilité du mouvement, Ann. Math. Stud. 17, Princeton Univ., 1947.

9 A. Ostrowsky and H. Schneider, Some theorems on the inertia of general matrices, *J. Math. Anal. Appl.* 4:72–84 (1962).
10 O. Taussky, A generalization of a theorem by Lyapunov, *J. Soc. Ind. Appl. Math.* 9:640–643 (1961).
11 J. Snyders and M. Zakai, On nonnegative solutions of the equation $AD + DA' = -C$, *SIAM J. Appl. Math.* 18:704–714 (1970).
12 H. K. Wimmer, Inertia theorems for matrices, controllability and linear vibrations, *Linear Algebra Appl.* 8:337–343 (1974).

ON A PARALLEL ALGORITHM FOR INHERENT SERIAL TECHNIQUES

by AVI LIN[28]

The present manuscript describes a new parallel algorithm to solve general linear recurrence systems with the dimension of N. Nowadays it is well recognized that one of the main bottlenecks in the realization of a parallel algorithm is the communication-time complexity. The important feature of the present algorithm is that it minimizes the broadcasting time (which can be shown to depend linearly on N).

In a little more detail this algorithm can be summarized as follows: Given the linear system of equations

$$\mathbf{A}\mathbf{y} = \mathbf{b} \tag{1}$$

for the vector y, where $\mathbf{A}$ is an $(m+n)\times n$ matrix with $a_{i,j} = 0$ for $j > m + i$, it is desired to solve for the last n components of y in terms of its first m components. Thus, $\mathbf{A}$ can be described as

$$\mathbf{A} \equiv \left[\mathbf{D}_{mn} \mid \mathbf{L}_n\right], \tag{2}$$

where $\mathbf{D}_{mn}$ is a general $m \times n$ matrix. We define also

the augmented $(m+1)\times n$ matrix $\mathbf{B}$ as

$$\mathbf{B} \equiv \left[\mathbf{b} \mid \mathbf{D}\right], \tag{3a}$$

[28] Computer Science Department, Technion–Israel Institute of Technology, Haifa, Israel, 32000.

the n-element vector $\mathbf{x}$ with the components

$$x_i \equiv y_{m+i}, \qquad 1 \leqslant i \leqslant n, \tag{3b}$$

the formal vector $\mathbf{d}$ as

$$\mathbf{d}^T \equiv (1, -y_1, -y_2, \ldots, -y_m). \tag{3c}$$

Thus the solution to Equation (1) can be carried out as follows: assume that $\mathbf{x} = \mathbf{Cd}$, where $\mathbf{C}$ in an $(m+1) \times n$ matrix, then find a $\mathbf{C}$ such that

$$\mathbf{Lx} = \mathbf{Bd}. \tag{3d}$$

This problem, which is denoted by L_DIAG$(m, n; *)$, and can be also presented as follows: transform the matrix $[\mathbf{B}_{m+1,n} | \mathbf{L}_n]$ into the matrix $[\mathbf{C}_{m+1,n} | \mathbf{I}_n]$, by applying row operations only, while using only $(*)$ processors, and where $\mathbf{I}_n$ is the $n \times n$ identity matrix. The main issue of the present paper is the following: given a P-processor machine, it is desired to solve the problem L_DIAG$(M, N; P)$ for large N by using all the available P processors, trying to minimize both their idling time and their communication time while keeping a good efficiency coefficient for any number P of processors. The following new algorithm is suggested:

Step 0: *Splitting.* Divide the matrix $[\mathbf{B} | \mathbf{L}]$ into P (not necessarily equal) horizontal pieces (strips). The rth strip has the width (number of rows or equations) of n_r, and its first row contains $m_r + 1$ nonzero elements, where

$$m_r = \sum_{j=1}^{r-1} n_j + M, \qquad r = 1, 2, \ldots, P. \tag{4}$$

Step 1: *Local elimination.* Assign to every strip r one (say the rth) processor element, and solve with this processor the problem L_DIAG$(n_r, m_r; 1)$. This step will be terminated by all the processors at the same time. This time is denoted by K.

Step 2: *Global elimination.* For $r = 1, 2, \ldots, P-1$ take the x results from the rth strip and substitute them into the remaining $N - n_{r+1}$ equations. The time needed for this step to be terminated is denoted by R.

The main problem here is to find such a partitioning n_i, $1 \leqslant i \leqslant P$, that the condition expressed in Step 1 will be fulfilled. Such a strategy is given by

the following lemma:

LEMMA 1. *If the time needed for all the processors to finish their tasks of local elimination is the same, then the following relations hold between the widths n_i of the various strips and K:*

$$n_i\left(M+1+\sum_{j=1}^{i-1} n_j\right)(\bar{\beta}+n_i)=2K, \qquad 1\leqslant i\leqslant P, \tag{5a}$$

$$\sum_{j=1}^{P} n_j = N \tag{5b}$$

where $\bar{\beta}$ is the ratio between the time for subtraction and the time for addition plus a multiplication.

This lemma presents a system of $P+1$ equations for $P+1$ unknowns which are all the n_i's and the time K. For comparison purposes, the "undimensional" time $\lambda \equiv K/[(M+1)N^2]$ will be used for the time of the first phase of the algorithm, and similarly $\sigma \equiv R/(MN^2)$ for the time of the second phase. An important feature of this splitting strategy is that for $M \geqslant 0$ and $\beta \geqslant 0$ the solution for the system of equations given in Lemma 1 yields

$$n_1 > n_2 > n_3 > \cdots > n_P.$$

Lemma 1 serves for finding the optimal strip division before the algorithm is executed. Its equations (5) lead to the following lemma:

LEMMA 2. *For $M/N \ll 1$ and $P/N \ll 1$, the local-elimination step of the present algorithm can be optimally applied if n_j, $2 \leqslant j \leqslant P$, are equal to one another and to $\sqrt{N(M+1)}$; the optimal distribution in the other cases ($M/N \sim 1$ and/or $P/N \sim 1$) is when all the equations are spread evenly over the P processors.*

Using these results, the speedup of this algorithm for $P \ll N \gg M$ is roughly

$$S(M,N;P)=\frac{1}{\left[1-\dfrac{P-1}{\sqrt{(M+1)N}}\right]\left[1-\dfrac{P-2}{P}\dfrac{P-1}{\sqrt{(M+1)N}}\right]} \tag{6}$$

when all the processors are used during the global elimination phase of the algorithm. Checking numerically the variation of the efficiency η with the

relative number of processors $\xi = (P-1)/\sqrt{N(M+1)}$, one finds that η attains its maximum value of 1 in all cases for $N \approx P^2/(M+1)$, and it attains a minimum value for all N's in the range of $0.3 \leqslant \xi \leqslant 0.5$. Unfortunately, the results obtained from Lemma 1 are poor compared to those obtained from Lemma 2, leading to the conclusion that the assumption given in Lemma 2 can be used only for $P/\sqrt{N(M+1)} < 0.3$.

The following lemma deals with the last relevant relations among N, P, and M.

LEMMA 3. *For $P \sim N$, Step 1 of the present algorithm is optimally applied for*

$$n_i \approx N/P, \qquad i = 1,2,\ldots,P. \tag{7}$$

For communication we assume the following simple model: when a processor element broadcasts u packets of information (say bytes or words), the broadcasting time T measured in units of t will be

$$T = a + bu, \tag{8}$$

where a is the "waking" time for the communication controller of the PE. With this model it can be shown that the minimum communication time complexity is $O(MN)$ and also the following result holds:

LEMMA 4. *The communication time that the present algorithm uses is optimal, and for the case where the equations are not redistributed in the global-elimination phase, the total time is*

$$T_{\text{total}} = (P-1)ar + (M+1)N^2\left[\frac{br}{N}\left(1 - \frac{n_P}{N}\right) + \sigma(N,P)\right] \tag{9}$$

where r is the ratio between a unit of communication time and a unit of a computation time.

The fact that the communication time in all other algorithms is larger than the communication time spent by the present algorithm by a factor of $\approx N/2$ makes this algorithm very attractive.

We now consider stability. One of the popular measures for the stability of a parallel algorithm in comparison with the stability of the corresponding serial algorithm is the ratio ρ of the number of computational operations of the parallel algorithm to that of the sequential one. Considering operations of

the type $(+\&*)$, for the case where $P \ll N$ or $M \ll N$, it can be shown that

$$\rho = 1 + (P-1)\left[1 - \frac{P-1}{\sqrt{N(M+1)}}\right]^2, \tag{10a}$$

while for the other cases

$$\rho = 1 + O\left(\frac{1}{P^3}\right). \tag{10b}$$

Thus the influence of the instability is bounded, and it attains its maximum value near the end of the region where Equation (10a) is applicable. Using some additional theorems, the following maximum value for ρ can be obtained:

$$\rho \approx \frac{P+1}{2}. \tag{11}$$

This result gives a realistic estimation for the stability of the present algorithm.

The general approach of the present algorithm can be used also to find the inverse of a triangular matrix in a parallel manner by an almost straightforward extension. Let us denote the problem of solving L_DIAG$(M, N; P)$ for Q RHS vectors by L_DIAG$(M, N, Q; P)$, with the understanding that L_DIAG$(M, N, 1; P) =$ L_DIAG$(M, N; P)$. Here the new $(M+Q)\times N$ matrix $\mathbf{R}$ is defined by taking the matrix $\mathbf{B}$ and appending to it, on its left, Q column vectors which are the Q RHS vectors $\mathbf{b}$ with opposite signs. Thus $\mathbf{R}$ has the following structure:

$$\mathbf{R} = \left[-\mathbf{b}_1 \mid -\mathbf{b}_2 \mid \cdots \mid -\mathbf{b}_Q \mid \mathbf{B}_{MN} \right]. \tag{12}$$

Let us first state the following equivalence lemma.

LEMMA 5. *A parallel solution of the problem* L_DIAG$(M, N, Q; P)$ *can be obtained by solving a problem of the type* L_DIAG$(M+Q, N; P)$ *with the present algorithm, where the matrix* $[\mathbf{B}\,|\mathbf{L}]$ *in the algorithm formulation has to be replaced by the matrix* $[\mathbf{R}\,|\mathbf{L}]$.

The general inverse of the matrix **A** of the type given by Equation (1) can now be defined by solving the problem L_DIAG($M, N, N; P$), where the RHS is composed of N vectors $\mathbf{b}_i$, $1 \leqslant i \leqslant N$, where $\mathbf{b}_i$ is a 0 vector except that in its ith entry it has a 1. This means that the exact ("pure") inverse of an L-matrix is obtained for $M = 0$. According to Lemma 5 this can be done by solving the problem L_DIAG($M + N, N; P$). The most interesting feature of the numerical results is that the value of Pn_1/N approaches an asymptotic value for quite small P. This fact is summarized in the following Lemma.

LEMMA 6. *The distribution of the equations for the optimal parallel solution of* L_DIAG($N, N, N; P$) *with the present algorithm is such that as* P/N *increases the value of* Pn_1/N *monotonically increases like* $1/(1 - 0.145P^{\Gamma})$ *with* $0 < \Gamma \ll 1$.

Numerical results show that the average value of Γ is 0.055... and increases very slowly with P. It now can be proven that the computational times spent in this case are $\lambda \approx 4(3 - \sqrt{2})/P^2$ and $\sigma \approx \frac{4}{3}(17 - 12\sqrt{2}) + (2 - \sqrt{2})/P$. The speedup of the computational time only is $\approx P/6(2 - \sqrt{2})$, and the efficiency is ≈ 0.285.

Another algorithm for the inverse, which is better than the previous one, is also presented:

Step 0: *Splitting.* Divide the matrix **R** into P horizontal strips. The rth strip has width of n_r, and its ith row contains $m_r + 2i - 1$ nonzero elements, where m_r is defined by Equation (4).

Step 1: *Local elimination.* Every strip is assigned to a processor which solves L_DIAG($n_r, m_r, n_r; 1$), taking into account the positions of the zero entries. This step will be designed so that all the processors will finish their work defined by this step at the same time. The time for termination of this step is denoted by K.

Step 2: *Global elimination.* For $r = 1, 2, \ldots, P - 1$ take the results from the rth strip, and substitute them into the remaining $N - m_{r+1}$ equations, with caution about zero substitution or "zero" subtraction. The time for this algorithm to be terminated is denoted by R.

The problem in implementing this algorithm is similar to the problem that was discussed for L DIAG($M, N; P$): how to partition the matrix **R** in the splitting step so that the condition expressed in Step 1 will be fulfilled. Such a strategy is given by the following lemma:

LEMMA 7. *If the time needed for all the processors to finish their tasks in the local-elimination phase is the same, then the following relations*

between the width of the ith strip, n_i and K holds:

$$n_i\left(3M+3+3\sum_{j=1}^{i-1} n_j + n_i\right)(\bar{\beta}+n_i) = 6K, \qquad 1 \leqslant i \leqslant P, \tag{13a}$$

subject to

$$\sum_{j=1}^{P} n_j = N. \tag{13b}$$

This lemma presents a system of $P+1$ equations for $P+1$ unknowns which are all the n_i's and the time K. R can be also approximated as follows:

LEMMA 8. *The time needed for the global substitution without dynamic redistribution of the rows among the processors is*

$$R = (M+1)N^2\sigma + N^3\tau, \tag{14a}$$

where

$$\tau \equiv \frac{1}{N^3}\sum_{j=1}^{P-1} n_j\left[\sum_{r=1}^{j} n_r n_{r+1} - \tfrac{1}{2}n_j n_{j+1}\right]. \tag{14b}$$

Like Lemma 1 for the previous algorithm, Lemma 7 in the present case dictates the splitting strategy of the system. This system is very similar to that presented by Lemma 1. However, it has quite different features, as shown by the following lemma:

LEMMA 9. *The set of reals $[n_i/n_1]$, where the set $[n_i]$ is given by Equation* (13a), *does not depend on N or P for small values of β and M.*

Using this lemma, one can easily prove the following approximation:

LEMMA 10. *The splitting of the system in phase 0 of the algorithm is such that*

$$\frac{n_i}{n_1} \approx \frac{C}{i^{\Gamma}}, \tag{15}$$

where $\frac{1}{3} < \Gamma < \frac{1}{2}$ and the value of Γ is close to $\frac{1}{3}$.

Redefining $\lambda \equiv K/N^3$, it can be shown that $\lambda = \frac{1}{6}(n_1/N)^3$, where the global-elimination phase and $M \sim 1$, τ is the undimensional time, as can be verified from Equation (14a). All the quantities λ, σ, and τ can be shown to decrease monotonically with increasing P. It can also be proven that

$$\tau = \frac{(1-\Gamma)^2}{2-3\Gamma} P^{-1}, \qquad \lambda = \frac{(1-\Gamma)^3}{6C^3} P^{3(\Gamma-1)}. \tag{16}$$

Similar results were obtained from numerical experiments. Using Equations (16) it can be shown also that to first order, the efficiency is constant and equal to $(1-\Gamma)^2/(2-3\Gamma) = \frac{4}{9}$.

The total communication time T_c that this algorithm uses is

$$\frac{T_c}{N^2} = (1-\Gamma)\left[1 - \frac{1-\Gamma}{2(1-2\Gamma)}\frac{1}{P}\right]$$

This amount of time is smaller than the $T_c/N^3 \approx 1$ in the other algorithms.

DIMENSIONS OF FACES GENERATED BY CERTAIN POSITIVE LINEAR OPERATORS

by RAPHAEL LOEWY[29]

Let R^n denote the vector space of all real n-tuples. A subset K of R^n is said to be a convex cone if it is convex and $\alpha x \in K$ for all $\alpha \geqslant 0$, $x \in K$. The cone K is pointed if $K \cap (-K) = \{0\}$, and proper if it is pointed and closed and spans R^n. All cones considered here are assumed to be proper.

A cone K induces a partial order, namely, for x, y in R^n define $x \leqslant y$ if and only if $y - x \in K$. A face F of K is a subcone of K such that $0 \leqslant y \leqslant x$ and $x \in F$ imply $y \in F$. Any one-dimensional face of K is called an extreme ray. Note that K must have at least n extreme rays. If $x \in K$, $x \neq 0$, and the ray through x is an extreme ray, we say x is an extremal of K. We let Ext K be the set of all extremals of K. It is well known that K is the convex hull of its extreme rays.

A cone K in R^n is said to be polyhedral if it has finitely many extreme rays, simplicial if it has exactly n extreme rays, and minimal if it has $n+1$

[29] Department of Mathematics, Technion–I.I.T., Haifa 32000, Israel.

extreme rays. A cone K is said to be decomposable if there exist two nonzero faces F_1 and F_2 such that $K = F_1 + F_2$ and span $F_1 \cap$ span $F_2 = \{0\}$. Otherwise K is said to be indecomposable.

Given two cones K_1 and K_2 in R^n and R^m, respectively, let $\pi(K_1, K_2) = \{A \in R^{m,n}: AK_1 \subset K_2\}$. If $K_1 = K_2 = K$, we write $\pi(K)$ for $\pi(K_1, K_2)$. Since $\pi(K_1, K_2)$ is itself a proper cone, the following problem is of great interest:

PROBLEM 1. Find the set Ext $\pi(K_1, K_2)$.

Generally, this seems to be a very difficult problem. One exception is the determination of the rank-one extremals of $\pi(K_1, K_2)$: see [5] and [7]. We describe now some results associated with Problem 1. We assume K_1 is not simplicial, for the case that K_1 is simplicial is obvious.

Fiedler and Pták [4] have characterized Ext $\pi(K_1, K_2)$ in case K_1 and K_2 are indecomposable minimal cones. Their interesting result is somewhat too complicated to be quoted in this short note. However, it follows from their analysis for this case that for some $r > 0$ the set of ranks of members of Ext $\pi(K_1, K_2)$ is $\{1,3,4,\ldots,r\}$.

Adin [1] introduced the concept of Gale pairs to study the set Ext $\pi(K_1, K_2)$ for K_1, K_2 polyhedral. He describes a procedure to determine Ext $\pi(K_1, K_2)$. In particular, he considers the case that K_1 is minimal and K_2 is an arbitrary polyhedral cone, and generalizes the result of Fiedler and Pták. Tam [10] also obtained results concerning Ext $\pi(K_1, K_2)$ in case K_1 is minimal, generalizing the results of [1] and [4].

Loewy and Schneider [7] characterized Ext $\pi(K_n)$, where $K_n = \{x \in R^n: x_n \geq (\sum_{i=1}^{n-1} x_i^2)^{1/2}\}$ and $n \geq 3$. They showed that the only extremals of $\pi(K_n)$ of rank greater than one are operators which map K_n onto itself.

The preceding results deal with some special cases of Problem 1. The following result gives sufficient conditions for $A \in \pi(K_1, K_2)$ to be an extremal.

THEOREM 1. *Let K_1 and K_2 be cones in R^n and R^m, respectively, and $A \in R^{m,n}$. Suppose the following conditions hold:*

(i) *K_1 is indecomposable.*
(ii) Ker $A = \{0\}$.
(iii) A Ext $K_1 \subseteq$ Ext K_2.

Then $A \in$ Ext $\pi(K_1, K_2)$.

This result was stated by Loewy and Schneider [7] for the case $K = K_1 = K_2$, but the same proof goes through for arbitrary K_1 and K_2. O'Brien [9] has

constructed an example of an indecomposable cone K and a nonsingular $A \in \operatorname{Ext} \pi(K)$ that does not take $\operatorname{Ext} K$ into itself.

In general, none of the hypotheses of Theorem 1 may be omitted. However, if $A \in \pi(K_1, K_2)$ and has rank r, it is of interest to obtain bounds on the dimension of $\phi(A)$, the face generated by A. So suppose now K_1 and K_2 are cones in R^n and R^m, respectively, and $A \in R^{m,n}$. Suppose that A, K_1, K_2 satisfy

(iii)′ $A \operatorname{Ext} K_1 \subseteq \operatorname{Ext} K_2 \cup \{0\}$;
(iv) K_1 and AK_1 are indecomposable cones;
(v) $\operatorname{rank} A = r$.

We ask

PROBLEM 2. When do we get the inequality

$$\dim \phi(A) \leqslant n + 1 - r? \tag{I}$$

Loewy [8] showed that (I) holds if K has $n+1$ or $n+2$ extreme rays. There exist examples where K_1 has $n+3$ extreme rays and (I) fails to hold. If the cone generated by any $n+2$ of the extreme rays of K_1 is indecomposable and the same holds for the cone generated by any $r+2$ of the extreme rays of AK_1, then (I) holds.

Proofs of unpublished results will appear elsewhere.

REFERENCES

1 R. Adin, Extreme positive operators on minimal and almost minimal cones, *Linear Algebra Appl.* 44:61–86 (1982).
2 G. P. Barker, The lattice of faces of a finite dimensional cone, *Linear Algebra Appl.* 7:71–82 (1973).
3 G. P. Barker, Theory of cones, *Linear Algebra Appl.* 39:263–291 (1981).
4 M. Fiedler and V. Pták, The rank of extreme positive operators on polyhedral cones, *Czechoslovak Math. J.* 28 (103):45–55 (1978).
5 E. Haynsworth, M. Fiedler, and V. Pták, Extreme operators on polyhedral cones, *Linear Algebra Appl.* 13:163–172 (1976).
6 R. Loewy and H. Schneider, Indecomposable cones, *Linear Algebra Appl.* 11:235–245 (1975).
7 R. Loewy and H. Schneider, Positive operators on the n-dimensional ice-cream cone, *J. Math. Anal. Appl.* 49:375–392 (1975).
8 R. Loewy, unpublished.
9 R. C. O'Brien, On extreme matrices and extreme vectors of cones in R^n, *Linear Algebra Appl.* 12:77–79 (1975).
10 B. S. Tam, Extreme positive operators on minimal cones, preprint.

HIGHER-DIMENSIONAL EUCLIDEAN AND HYPERBOLIC MATRIX SPACES

by BINYAMIN SCHWARZ[30]* and ABRAHAM ZAKS[30]

1. *Introduction*

Complex $n \times n$ matrices are denoted by capital Latin letters $P = (p_{ik})_1^n$. Complex $n \times mn$ $(n, m \geqslant 2)$ matrices are denoted by capital script letters and are often written in block form $\mathscr{P} = (P_1 \ \cdots \ P_m)$, where each P_i is an $n \times n$ matrix. The set of all $n \times mn$ matrices $\mathscr{P}$ of rank $\rho(\mathscr{P}) = n$ is called $C_0(m, n)$, and only such full-rank matrices will be used. Two such matrices $\mathscr{P}$ and $\tilde{\mathscr{P}}$ are (row- or left-) equivalent if there exists an invertible $n \times n$ matrix R such that $\tilde{\mathscr{P}} = (\tilde{\mathscr{P}}_1 \ \cdots \ \tilde{P}_m) = (RP_1 \ \cdots \ RP_m) = R\mathscr{P}$, $|R| \neq 0$. The corresponding equivalence classes are the points of $P = P_{m-1}(M_n(C))$, the *m-dimensional projective space over the complex $n \times n$ matrices*. The point corresponding to the matrix $\mathscr{P}$ is denoted by $\mathbf{P}$, and we write $\mathbf{P} = f(\mathscr{P})$, $\mathscr{P} \in f^{-1}[\mathbf{P}]$ and call f the *standard map* from $C_0(m, n)$ to P. The topology of P is the quotient topology relative to f and the usual topology of $C_0(m, n)$. *P is a connected and compact Hausdorff space whose topology has a countable base* [3].

In [4] we considered several metrics for the projective matrix line $P_1(M_n(C))$. To define these metrics also for the higher-dimensional space $P = P_{m-1}(M_n(M_n(C))$, $m > 2$, we embed P into the line P' of matrices of appropriate dimension $P' = P_1(M_{mn-n}(C))$. We add primes to the letters used for matrices and points of this space. P' is thus an $(mn - n) \times (mn - n)$ matrix; $(mn - n) \times 2(mn - n)$ matrices are written in block form $\mathscr{P}' = (P_1' \ P_2')$. The set of all full-rank matrices of this size is called $C_0(2, mn - n)$. $C_0(m, n)$ is mapped into $C_0(2, mn - n)$ in the following way. For every $\mathscr{P} = (P_1 \ \cdots P_m) \in C_0(m, n)$ we define two $(mn - n) \times (mn - n)$ matrices

$$P_1' = \begin{pmatrix} P_1 & \cdots & P_{m-1} \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}, \qquad P_2' = \begin{pmatrix} P_m & 0 & \cdots & 0 \\ 0 & I & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & I \end{pmatrix},$$

and set $\mathscr{P}' = (P_1' \ P_2') = g(\mathscr{P})$. This map induces a map from $P =$

[30] Department of Mathematics, Technion–Israel Institute of Technology, Haifa 32000, Israel.

$P_{m-1}(M_n(C))$ into $P' = P_1(M_{mn-n}(C))$:

$$\mathbf{P}' = f'(\mathscr{P}') = f'(g(\mathscr{P})) = f'(g(f^{-1}[\mathbf{P}])) = h(\mathbf{P}).$$

The map h from P to $h(P) \subset P'$ is a homeomorphism [6].

2. *The Euclidean Geometry*

The set P_f of finite points of P is defined by $P_f = \{\mathbf{P}: \mathscr{P} = (P_1 \ \dots \ P_m) \in f^{-1}[\mathbf{P}], \ |P_m| \neq 0\}$. For any finite point $\mathbf{P}$ we use, in this section, its uniquely given canonical matrix $\mathscr{P} = (P_1 \dots P_{m-1} I) \in f^{-1}[\mathbf{P}]$. The map of a canonical matrix is again canonical: $\mathscr{P}' = g(\mathscr{P}) = (P' \ I')$, hence $h(P_f) \subset (P')_f$. The Euclidean distance of $\mathbf{P} = f(P_1 \ \dots \ P_{m-1} \ I)$ and $\mathbf{Q} = f(Q_1 \ \dots \ Q_{m-1} \ I)$ is defined by $d(\mathbf{P},\mathbf{Q}) = \|\sum_{i=1}^{m-1}(P_i - Q_i)(P_i - Q_i)^*\|^{1/2}$, where $\| \ \|$ denotes the spectral norm.

THEOREM 1. *Let $\mathbf{P}$ and $\mathbf{Q}$ be points of $P_f \subset P = P_{m-1}(M_n(C))$, and let $\mathbf{P}' = h(\mathbf{P})$ and $\mathbf{Q}' = h(\mathbf{Q})$ be their images under the map h. Then $\mathbf{P}'$ and $\mathbf{Q}'$ belong to $(P')_f \subset P' = P_1(M_{mn-n}(C))$ and $d(\mathbf{P},\mathbf{Q}) = d(\mathbf{P}',\mathbf{Q}')$. The function $d(\mathbf{P},\mathbf{Q})$ defines a metric for P_f.*

The finite part P_f of $P_{m-1}(M_n(C))$ with this metric may be called $E = E^{m-1}(M_n(C))$, the *$(m-1)$-dimensional Euclidean space over the complex $n \times n$ matrices*, and we use the notation $\mathbf{P} = (P_1, \dots, P_{m-1})$ instead of $\mathbf{P} = f(P_1 \ \dots \ P_{m-1} \ I)$. Linear dependence and linear subspaces in $P_{m-1}(M_n(C))$ were considered in [5]. The hyperplane π in E is given by

$$P_1 A_1^* + \cdots + P_{m-1} A_{m-1}^* = A^*, \tag{1}$$

where $\mathscr{A}_0 = (A_1 \ \dots \ A_{m-1})$ is of rank n. As we may replace $\mathscr{A}_0$ and A by $R\mathscr{A}_0$ and RA, $|R| \neq 0$, we may assume that

$$A_1 A_1^* + \cdots + A_{m-1} A_{m-1}^* = I. \tag{2}$$

THEOREM 2. *Let the hyperplane π of $E = E^{m-1}(M_n(C))$ be given in its normal form, i.e. by Eqs.* (1) *and* (2). *Let $\mathbf{Q} = (Q_1, \dots, Q_{m-1})$ be an arbitrary point of E. Then $d(\mathbf{Q}, \pi) = \|Q_1 A_1^* + \cdots + Q_{m-1} A_{m-1}^* - A^*\|$.*

3. *The Hyperbolic Geometry*

Let $\mathbf{O} = f(0 \ \cdots \ 0 \ I)$ be the origin of $P = P_{m-1}(M_n(C))$, and let $\Delta = \{\mathbf{P}: d\mathbf{O},\mathbf{P}) < 1\}$ be the unit ball of P. We use now the subset J of

$C_0(m,n)$:

$$J = \left\{ \mathscr{P}: \mathscr{P} = (P_1 \ \dots \ P_m),\ P_m P_m^* - \sum_{i=1}^{m-1} P_i P_i^* = I \right\}$$

and have $\Delta = \{\mathbf{P}: \mathscr{P} \in f_{-1}[\mathbf{P}],\ \mathscr{P} \in J\}$. For any pair $\mathbf{P},\mathbf{Q}$ of points in Δ let $\mathscr{P} = (P_1 \ \cdots \ P_m) \in f^{-1}[\mathbf{P}]$, $\mathscr{Q} = (Q_1,\dots,Q_m) \in f^{-1}[\mathbf{Q}]$, $\mathscr{P},\mathscr{Q} \in J$, and define $H(\mathscr{P},\mathscr{Q}) = P_m Q_m - \sum_{i=1}^{m-1} P_i Q_i$. We set $\rho(\mathbf{P},\mathbf{Q}) = \|H(\mathscr{P},\mathscr{Q})H^*(\mathscr{P},\mathscr{Q}) - I\|^{1/2}$ and define the pseudochordal distance $\psi(\mathbf{P},\mathbf{Q}) = \rho(\mathbf{P},\mathbf{Q})/[1+\rho^2(P,Q)]^{1/2}$ (cf. [2]).

THEOREM 3. *Let* $\mathbf{P}$ *and* $\mathbf{Q}$ *be points of the unit ball* Δ *of* $P = P_{m-1}(M_n(C))$, *and let* $\mathbf{P}' = h(\mathbf{P})$ *and* $\mathbf{Q}' = h(\mathbf{Q})$ *be their images under the map* h. *Then* $\mathbf{P}'$ *and* $\mathbf{Q}'$ *belong to the unit ball* $\Delta' \subset P' = P_1(M_{mn-1}(C))$, *and* $\psi(\mathbf{P},\mathbf{Q}) = \psi(\mathbf{P}',\mathbf{Q}')$. *The function* $\psi(\mathbf{P},\mathbf{Q})$ *defines a metric for* Δ.

The pseudochordal distance ψ leads to the non-Euclidean (hyperbolic) distance E_n.

THEOREM 4. *The function*

$$E_n(\mathbf{P},\mathbf{Q}) = \tfrac{1}{2}\log\frac{1+\psi(\mathbf{P},\mathbf{Q})}{1-\psi(\mathbf{P},\mathbf{Q})} = \operatorname{arcsinh}\rho(\mathbf{P},\mathbf{Q})$$

defines an intrinsic metric for the unit ball Δ *of* $P = P_{m-1}(M_n(C))$.

In the real scalar case, i.e. for the ordinary real $(m-1)$-dimensional projective space $P_{m-1}(R)$, this metric reduces to the metric of the hyperboloid model [1].

A detailed version of these results will be given in [7].

This research was supported by the fund for the promotion of research at the Technion.

REFERENCES

1 A. F. Beardon, *The Geometry of Discrete Groups*, Springer, New York, 1983.
2 C. Carathéodory, *Theory of Functions of a Complex Variable*, Vol. 1, Chelsea, New York, 1954.
3 B. Schwarz and A. Zaks, Matrix Möbius transformations, *Comm. Algebra* 9 (19):1913–1968 (1981).
4 B. Schwarz and A. Zaks, Geometries of the projective matrix space, *J. Algebra* 95 (1):263–307 (1985).
5 B. Schwarz and A. Zaks, Geometry of matrix differential systems, *J. Math. Anal. Appl.*, 112:165–177 (1985).

6 B. Schwarz and A. Zaks, On the embeddings of projective spaces in lines, *Linear and Multilinear Algebra* 18:319–336 (1985).
7 B. Schwarz and A. Zaks, Higher-dimensional Euclidean and hyperbolic matrix spaces, *J. Analyse Math.*, to appear.

FROM COMPLEX NUMBERS TO COMPLEX MATRICES ALONG THE PROJECTIVE LINE

by BINYAMIN SCHWARZ[31] and ABRAHAM ZAKS[31*]

In the theory of functions one compares the Euclidean plane **E** and the Riemann sphere **S** by the stereographic projection. The Möbius transformations $w=(zc+d)^{-1}(za+b)$, $ad-bc=1$, are used in the study of the unit disk **D** and the upper half plane **H**. These parts of the classical theory lead to the use of projective geometry. The projective line **P** consists of equivalence classes of nonzero pairs $(z_1, z_2) \approx (rz_1, rz_2)$, where $z=z_1/z_2$. It results that the Möbius transformation is a linear mpa on **P** (where $w=w_1/w_2$):

$$w_1 = z_1 a + z_2 b, \qquad w_2 = z_1 c + z_2 d. \tag{1}$$

The complex numbers correspond to the finite points $(z,1)$ of **P**. The "north pole" $N=(1,0)$ represents the point at infinity.

By properly choosing $r=e^{i\theta}$ we may pick a representative for (z_1, z_2) so that z_2 is real. We may further choose r so that z_2 is real and $z_1 z_1^* + z_2 z_2^* = 1$, where z^* denotes the complex conjugate of z. If $z_1 = x+iy$ and $z_2 = z$, where x, y, z are real numbers, then these points lie on the unit sphere **B** in R^3, $x^2+y^2+z^2=1$. On **B**, (z_1, z_2) and $(-z_1, -z_2)$ correspond to the same point on **P**, and the equator $(z,0)$ corresponds to the north pole.

To obtain unique representatives, we now consider the points $(z_2^* z_1, z_2^* z_2)$ in R^3 that lie on the Riemann sphere **S** of radius $\frac{1}{2}$, centered at $(0,0,\frac{1}{2})$. Thus **S** may serve to represent **P**.

Antipodal points in **S** correspond to perpendicular points (z_1, z_2) and (z_3, z_4) in **B**. There results a 2×2 unitary matrix **Z** whose rows are these points (see Figure 1).

To study the points (x, y) of the Euclidean plane **E** we identify them with the finite points of **P**, that is, $(z,1)$. In R^3 these correspond to the points $(x, y, 1)$. The projective equivalence induces the map

$$(z,1) \in \mathbf{E} \to (uz, u) \in \mathbf{B} \to (u^2 z, u^2) \in \mathbf{S}. \tag{2}$$

[31] Department of Mathematics, Technion, Haifa 32000, Israel. This research was supported by the fund for the promotion of research at the Technion.

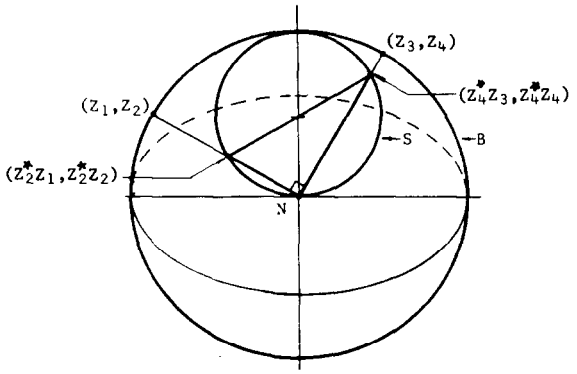


FIG. 1.

If **Z** and **W** are unitary matrices corresponding to two points (and their antipodes), then the chordal distance is derived from $\mathbf{R} = \mathbf{Z}\mathbf{W}^{-1}$:

$$\|R_{21}\| = \|R_{12}\| = d_s(z, w) = \left\|(1 + zz^*)^{-1/2}(w - z)(1 + w^*w)^{-1/2}\right\|, \quad (3)$$

while the intrinsic distance is obtained using arcsin. The unitary matrices represent the distance-preserving motions of **S**.

For the Euclidean plane **E**, the point at infinity is a fixed point. To a point z in **E** there corresponds a matrix **Z** whose rows are (1 0) and (z 1) (**Z** is a Euclidean motion). If w is another point in **E**, and **W** the corresponding matrix, then $\mathbf{R} = \mathbf{Z}\mathbf{W}^{-1}$ gives rise to the Euclidean distance $\|R_{21}\| = d_E(z, w) = \|z - w\|$. The Euclidean matrices preserve the Euclidean distance.

The stereographic projection of **E** into **S** may be derived from the fact that for finite points $(z_1, z_2) \in \mathbf{P}$ we have $(z_2^{-1}z_1, 1) \in \mathbf{E}$ and $(zz_1, zz_2) \in \mathbf{B}$, where $zz^* = (z_1z_1^* + z_2z_2^*)^{-1}$. Note that usually the stereographic projection maps $\mathbf{S} - \{N\}$ (or $\mathbf{B} - \{N\}$) into $\mathbf{E}_0 = \{(z, 0)\}$, while here the map is into $\mathbf{E} = \{(z, 1)\}$ (see Figure 2).

To investigate the unit disk **D**, we note that $(z_1, z_2) \in \mathbf{D}$ if $z_2z_2^* - z_1z_1^* > 0$. We may normalize by choosing a representative for which $z_2z_2^* - z_1z_1^* = 1$. Hence, if $z_2 = z$ and $z_1 = x + iy$ (with x, y, z real, $z > 1$), then $z^2 - x^2 - y^2 = 1$. The upper sheet of this hyperboloid of two sheets represents **D** in R^3. The points on $x^2 + y^2 - z^2 = 1$, $z > 0$, represent the points outside of **D**.

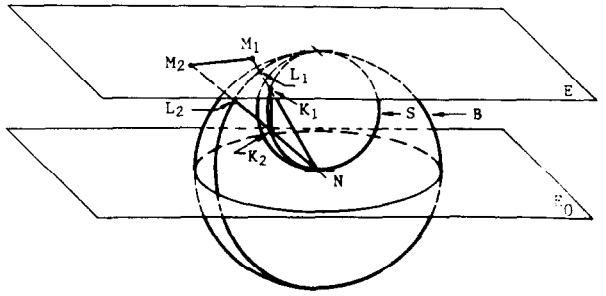


FIG. 2.

If $(z,1) \in \mathbf{D}$, then the inverse point is $(1, z^*)$. We may normalize them as

$$(z,1) \approx \left((1-zz^*)^{-1/2}z, (1-zz^*)^{-1/2}\right) = (z_3, z_4) \in \mathbf{D}, \tag{4}$$

$$(1,z^*) \approx \left((1-z^*z)^{-1/2}, (1-z^*z)^{-1/2}z^*\right) = (z_1, z_2). \tag{5}$$

Let $\mathbf{J}$ be the 2×2 matrix whose rows are $(1\ \ 0)$ and $(0\ \ -1)$, and let $\mathbf{Z}$ be the matrix whose rows are $(z_1\ \ z_2)$ and $(z_3\ \ z_4)$. Then $\mathbf{J} = \mathbf{ZJZ}^*$. We say that $\mathbf{Z}$ is $\mathbf{J}$-unitary. If $\mathbf{W}$ is another pair of inverse points inside and outside $\mathbf{D}$, then $\mathbf{R} = \mathbf{ZW}^{-1}$ gives rise to the non-Euclidean (hyperbolic) distance,

$$\|R_{21}\| = \|R_{12}\| = d_h(z,w) = \left\|(1-zz^*)^{-1/2}(z-w)(1-w^*w)^{-1/2}\right\|. \tag{6}$$

The intrinsic metric is achieved using arcsinh. The motions that preserve these distances are given by $\mathbf{J}$-unitary matrices.

The maps that carry the unit disk into itself are necessarily contractions.

The relations of the Riemann sphere, the Euclidean plane, and the unit disk are described in Figure 3.

The upper half plane $\mathbf{H}$ consists of the points $(z,1)$ for which $-i(z-z^*) > 0$. The conjugate point in the lower half plane is $(z^*,1)$. We choose a representative for the point on $\mathbf{H}$ that is $(z_1, z_2) = r(z,1)$, so that $-i(z_1z_2^* - z_2z_1^*) = 1$ (by choosing r so that $rr^* = [-i(z-z^*)]^{-1}$). In a similar way we choose the representative for $(z^*,1) \approx (z_3, z_4)$. We let $\mathbf{Z}$ denote the matrix whose rows are $(z_1\ \ z_2)$ and $(z_3\ \ z_4)$. Let $\mathbf{K}$ denote the matrix whose rows are $(0\ \ -i)$ and$(i\ \ 0)$. Then $\mathbf{ZKZ}^* = \mathbf{J}$.

We term $\mathbf{Z}$ a $\mathbf{K}$-unitary matrix if $\mathbf{ZKZ}^* = \mathbf{K}$. If $\mathbf{W}$ is a matrix representing another point in $\mathbf{H}$ and its conjugate, then $\mathbf{R} = \mathbf{ZW}^{-1}$ gives rise to the

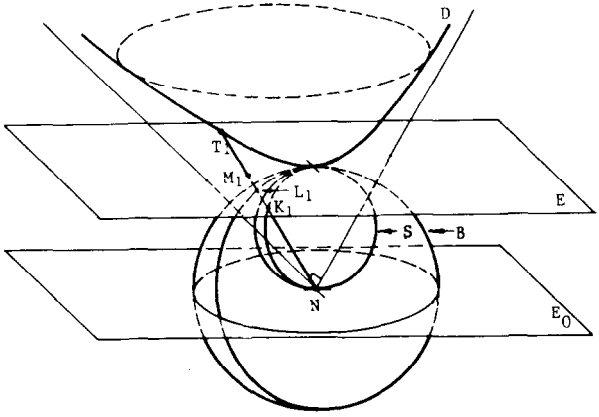


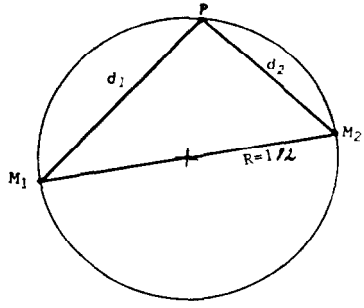
FIG. 3.

Poincaré metric (using arcsinh):

$$\|R_{21}\| = \|R_{12}\| = d_p(z, w) = \left\|(z - z^*)^{-1/2}(z - w)(w - w^*)^{-1/2}\right\|. \quad (7)$$

The metric-preserving motions are induced by **K**-unitary matrices. Since a unitary matrix **U** exists such that $\mathbf{UJU}^* = \mathbf{K}$, then $\mathbf{ZKZ}^* = \mathbf{J}$ whenever **UZ** is a **K**-unitary matrix.

The advantage of this approach using the projective-line model lies in the fact that it is possible to extend the results from the complex numbers to the matrices with complex entries, with the obvious changes, e.g., z^* denotes the



$$d_1^2 + d_2^2 = 1.$$

FIG. 4.

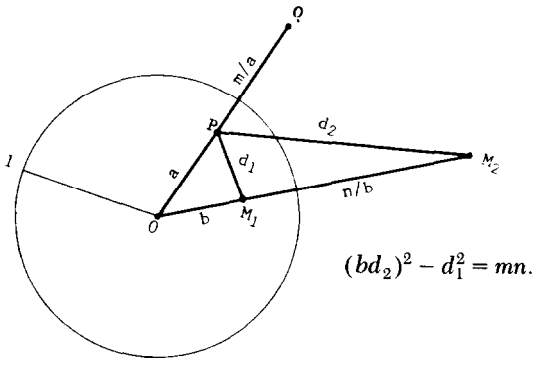


FIG. 5.

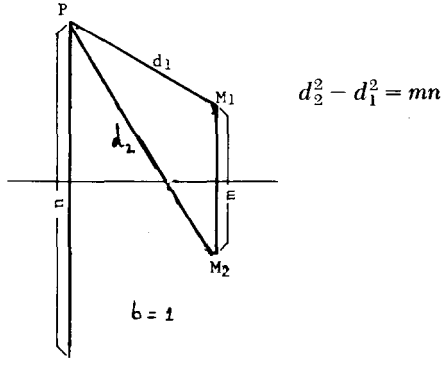


FIG. 6.

conjugate transpose of z, and $\| \ \|$ is the spectral norm. Further generalizations may be derived to general rings of operators. Furthermore, by allowing a change in the ring of scalars, we are able to study similar properties of higher-dimensional spaces by embedding them in a suitable projective line.

An interesting difference occurs in the study of the circles. In the matrix case it is necessary to consider *both centers*: in **S** the pair of antipodal points, in **D** the pair of inverse points, and in **H** the pair of conjugate points. We have to impose some equalities, in view of the fact that the circles have two centers and therefore two radii. These equalities hold naturally in the scalar case (see Figures 4, 5, 6).

REFERENCES

1 B. Schwarz and A. Zaks, Matrix Möbius transformations, *Comm. Algebra* 9:1913–1968 (1981).
2 B. Schwarz and A. Zaks, Geometries of the projective matrix space, *J. Algebra* 95:263–307 (1985).
3 B. Schwarz and A. Zaks, Geometries of the projective matrix space II, *J. Algebra* to appear.

ON THE UNIQUENESS OF THE LYAPUNOV SCALING FACTORS

by DAFNA SHASHA[32*] and ABRAHAM BERMAN[32,33]

A matrix $A \in R^{n,n}$ is (positive) *stable* if all its eigenvalues lie in the open right half plane. By Lyapunov's theorem [5] A is stable if and only if there exists a matrix $H > 0$ such that

$$AH + HA^T > 0, \tag{1}$$

where $X > 0$ means that X is positive definite. A real matrix A is *Lyapunov diagonally stable* if the matrix H in (1) can be chosen to be diagonal, i.e., if there exists a positive diagonal matrix D such that

$$AD + DA^T > 0. \tag{2}$$

A is *Lyapunov diagonally semistable* if there exists a positive diagonal matrix D such that

$$AD + DA^T \geqslant 0, \tag{3}$$

where $X \geqslant 0$ means that X is positive semidefinite.

Lyapunov diagonal stability and semistability play an important role in some problems in ecology and economics; see e.g. [1] and the references there.

The matrix D in (2) or (3) is called a *Lyapunov scaling factor* of A in [3], where the question of its uniqueness is studied. A related question is studied in [4].

[32] Department of Mathematics, Technion–Israel Institute of Technology, Haifa 32000, Israel.

[33] Research supported by Technion VPR Grant No. 100-641.

Clearly, if D is a Lyapunov scaling factor of A, then so is kD for every positive scalar k. Thus by saying that A has a *unique Lyapunov scaling factor* we mean uniqueness up to multiplication by a scalar. It is also clear, by continuity considerations, that a Lyapunov diagonally stable matrix does not have a unique Lyapunov scaling factor. Motivated by this observation, Hershkowitz and Schneider [3] defined A to be *Lyapunov diagonally near-stable* if it is Lyapunov diagonally semistable but not Lyapunov diagonally stable, and suggested the following conjecture:

CONJECTURE [3, Conjecture 6.32]. *If $A \in R^{n,n}$ is a Lyapunov diagonally near-stable, irreducible P-matrix (i.e., a matrix all of whose principal minors are positive), then A has a unique Lyapunov scaling factor.*

In this paper we settle the conjecture by giving a counterexample when $n = 4$ and proving it for $n \leqslant 3$.

A COUNTEREXAMPLE. Let

$$A = \begin{pmatrix} 1 & 0 & 2 & 4 \\ 2 & 1 & 0 & 4 \\ 0 & 2 & 1 & 4 \\ 4 & 4 & 4 & 20 \end{pmatrix}.$$

A is an irreducible P-matrix which is not stable and thus not Lyapunov diagonally stable. Let

$$D = \operatorname{diag}\{1,1,1,d\}.$$

Then

$$AD + DA^T \geqslant 0 \qquad \text{for} \quad \frac{3-\sqrt{5}}{2} \leqslant d \leqslant \frac{3+\sqrt{5}}{2}.$$

THEOREM 1. *Let $A \in R^{n,n}$ be a Lyapunov diagonally semistable matrix, and let D be a Lyapunov scaling factor of A. Then:*

(i) *A is Lyapunov diagonally near-stable if and only if there exists a nonzero positive semidefinite matrix B such that*

$$(BA)_{ii} = 0, \qquad i = 1, \ldots, n.$$

(ii) *B, A, and D satisfy*

$$B(AD + DA^T) = 0,$$

so

$$\operatorname{rank}(B) + \operatorname{rank}(AD + DA^T) \leqslant n.$$

THEOREM 2. *Let $A \in R^{n,n}$ be a Lyapunov diagonally near-stable P-matrix, and let D be a Lyapunov scaling factor of A. Then*

$$1 \leqslant \operatorname{rank}(AD + DA^T) \leqslant n - 2.$$

COROLLARY. *If $A \in R^{3,3}$ is a Lyapunov diagonally near-stable P-matrix, and $D > 0$ is a Lyapunov scaling factor of A, then*

$$\operatorname{rank}(AD + DA^T) = 1$$

THEOREM 3. *A 3×3 Lyapunov diagonally near-stable P-matrix has a unique Lyapunov scaling factor.*

Observe that irreducibility is not mentioned in Theorem 3. This is not surprising, since (e.g. [2]) every reducible 3×3 P-matrix is Lyapunov diagonally stable.

Proofs and examples are given in [6].

REFERENCES

1 G. P. Barker, A. Berman, and R. J. Plemmons, Positive diagonal solutions to the Lyapunov equations, *Linear and Multilinear Algebra* 5:249–266 (1978).

2 A. Berman and D. Hershkowitz, Matrix diagonal stability and its implications, *SIAM J. Algebraic Discrete Methods* 4:377–382 (1983).

3 D. Hershkowitz and H. Schneider, Scalings of vector spaces and the uniqueness of Lyapunov scaling factors, *Linear Multilinear Algebra* 17:203–226 (1985).

4 ———, Semistability factors and semifactors, *Contemp Math.*, 47:203–216 (1985).

5 A. Lyapunov, Problème générale de la stabilité du mouvement, *Comm Math. Soc. Kharkow*, 1892; republished in *Ann. Math. Stud.* 17, Princeton U.P., 1947.

6 D. Shasha and A. Berman, On the uniqueness of the Lyapunov scaling factors, *Linear Algebra Appl.*, to appear.

THE RESOLVENT CONDITION AND UNIFORM POWER-BOUNDEDNESS

by EITAN TADMOR[34]

Let L be an operator with uniformly bounded powers:

$$\|L^k\| \leqslant M_P, \qquad k = 1,2,\ldots. \tag{P}$$

Using the geometric expansion for the resolvent of such an operator, $(zI - L)^{-1}$, it follows that

$$\left\|(zI - L)^{-1}\right\| \leqq \frac{M_R}{|z| - 1} \qquad \text{for all} \quad |z| > 1, \tag{R}$$

with constant $M_R = M_P$.

In this talk we discuss the *inverse implication* of the above, namely, the power-boundedness of operators which satisfy the *resolvent condition* (R).

We begin with the finite-dimensional case, considering *families of matrices*. Thus, suppose L is given as a direct sum of finite-dimensional operators, their dimension being *uniformly* bounded, say $\leqslant N$. Then (R) $\Rightarrow$ (P) is in fact just one of the four implications contained in the Kreiss matrix theorem [6] which was subsequently treated by many authors, including [2–4], [7], [10–12], [14]. A simple derivation of this, which led to a power estimate sharper than the previous ones, was given at [15], asserting

$$\|L^k\| \leqslant \text{const}_R \cdot N, \qquad k = 1,2,\ldots,$$

with the linear dependence on the dimension N being the best possible [8].

Turning to the infinite-dimensional case, we first note—using an argument due to Sz.-Nagy [13]—that *compact* operators satisfying (R) are necessarily similar to contractions:

$$\|TLT^{-1}\| \leqslant 1,$$

and hence are power-bounded with $M_p = \|T\| \cdot \|T^{-1}\|$. (The existence of such similarity in the finite-dimensional setup was proved in [6], [11].)

[34] School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel-Aviv University, Tel-Aviv 69978, Israel.

Noncompact counterexamples of Foguel [1] and Halmos [5] indicate that the powers of general operators satisfying (R) may grow. How fast is the power growth permitted by the resolvent condition? An easy application of the Cauchy integral formula yields a *linear* upper bound:

$$\|L^n\| \leqslant e \cdot M_R \cdot n, \qquad n = 1,2,\ldots.$$

In some cases this estimate can be improved on the basis of the following

LEMMA [16]. *Suppose L satisfies the resolvent condition* (R). *Let d_n denote the following minimax quantity*:

$$d_n = \min_{\zeta} \max_{\eta} |\zeta - \eta| \cdot \left\| \left(\left(1 + \frac{1}{n} \right) e^{i\eta} - L \right)^{-1} \right\|.$$

Then the following estimate holds:

$$\|L^n\| \leqslant \text{const}_R \cdot d_n \log n, \qquad n = 2,3,\ldots.$$

The last result yields a *logarithmic power growth* provided the spectrum of L is not "too dense" in the neighborhood of the unit circle. One such case is the dissipative case, where instead of (R) we have the stronger *dissipativity condition*

$$\|(zI - L)^{-1}\| \leqslant \frac{M_D}{|z - 1|} \qquad \text{for all} \quad |z| > 1. \tag{D}$$

S. Friedland (private communication) has given an alternative proof of the logarithmic growth in this case. The same estimate applies if there is a *finite* number of simple poles on the unit circle. Finally, we give a counterexample satisfying (R), [9], with an *unbounded* number of such poles and with a logarithmic power growth.

REFERENCES

1 S. R. Foguel, A counterexample to a problem of Sz.-Nagy, *Proc. Amer. Math. Soc.* 15:788–790 (1964).

2 S. Friedland, A generalization of the Kreiss matrix theorem, *SIAM J. Math. Anal.* 12:826–832 (1983).

3 M. Gorelick and H. Kranzer, An extension of the Kreiss stability theorem to families of matrices of unbounded order, *Linear Algebra Appl.* 14:237–256 (1976).
4 E. Görlich and D. Pontzen, Alpha well-posedness, Alpha stability and the matrix theorems of H. O. Kreiss, *Numer. Math.* 46:131–146 (1985).
5 R. Halmos, On Foguel's answer to Nagy's question, *Proc. Amer. Math. Soc.* 15:791–793 (1964).
6 H.-O. Kreiss, Über die Stabilitätdefinition für Differenzengleichunger die Differentialgleichungen approximieren, *BIT* 2:153–181 (1962).
7 G. Laptev, Conditions for the uniform well-posedness of the Cauchy problem for systems of equations, *Soviet Math. Dokl.* 15:65–69 (1975).
8 R. J. LeVeque and L. N. Trefethen, On the resolvent condition in the Kreiss matrix theorem, *BIT*, to appear.
9 C. A. McCarthy and J. Schwartz, On the norm of finite Boolean algebra of projections and applications to theorems of Kreiss and Morton, *Comm. Pure Appl. Math.* 18:191–201 (1965).
10 J. H. Miller, On power-bounded operators and operators satisfying a resolvent condition, *Numer. Math.* 10:389–396 (1967).
11 J. H. Miller and G. Strang, Matrix theorems for partial differential and difference equations, *Numer. Math.* 10:113–123 (1966).
12 K. W. Morton, On a matrix theorem due to H.-O. Kreiss, *Comm. Pure Appl. Math.* 17:375–379 (1964).
13 B. Sz.-Nagy, Completely continuous operators with uniformly bounded iterates, *Magyar Tud. Akad. Kutato Int. Köl* 4:89–93 (1959).
14 R. D. Richtmyer and K. W. Morton, *Difference Methods for Initial-Value Problems*, 2nd ed., Interscience, New York, 1967.
15 E. Tadmor, The equivalence of L_2-stability, the resolvent condition and strict H-stability, *Linear Algebra Appl.* 41:151–159 (1981).
16 E. Tadmor, The logarithmic growth of operators satisfying the resolvent condition, preprint.

AN EFFICIENT PRECONDITIONING ALGORITHM AND ITS ANALYSIS

by M. TISMENETSKY[35] AND I. EFRAT[35*]

Introduction

The purpose of this work is to suggest and analyze a new preconditioning for solving sparse linear systems, which is readily vectorized and very efficient for matrices arising from a finite-difference discretization of partial

[35]IBM Scientific Center, Technion City, Haifa 32000, Israel.

differential equations of elliptic type. The algorithm is in fact, a variant of the incomplete-block-factorization technique accelerated by biconjugate gradients or other acceleration methods. It turns out that under practical conditions the algorithm does not break down and gives a significant improvement in the condition number and the resulting convergence rate.

Consider the equation

$$A\mathbf{x} = \mathbf{b}, \qquad A \in \mathbb{R}^{N\times N}, \tag{1}$$

with a sparse matrix A partitioned into block form. It is well known that iterative solution techniques are sensitive to the matrix spectral condition number $K(A)$. To reduce the condition number, a premultiplication of (1) by a matrix $\tilde{A}^{-1}$, usually called "preconditioning," is often exploited. The matrix $\tilde{A}$, termed the "preconditioner," is expected to be "close" to A, so that $K(\tilde{A}^{-1}A) \ll K(A)$. $\tilde{A}$ may not be available explicitly, however the number of arithmetic operations needed to perform the preconditioning must be of order N. Recent efficient preconditioning techniques (see, for instance, [1–5]) are responsible for dramatic improvements in the efficiency of iterative methods such as Conjugate or Biconjugate Gradients.

In this work we develop a preconditioning algorithm for sparse block matrices. For brevity, however, only the block-tridiagonal case with banded blocks will be considered in this synopsis. More precisely, in (1) let $A = [A_{ij}]_{i,j=1}^{m}$, where $A_{ij} \in \mathbb{R}^n$ are sparse and $A_{ij} = 0$ for $|i-j| > 1$, $i, j = 1, 2, \ldots, m$. Only such matrices are considered in the sequel. In presenting the algorithm based on an LU decomposition of A, we adopt the following notation. Given a vector $\mathbf{c}$, diag[$\mathbf{c}$] stands for the diagonal matrix formed from $\mathbf{c}$. The notation rs(C) is used for the vector of row sums of C, while inv(C) denotes an approximate inverse of the matrix C. Also, $C^{(p)}$ stands for a p-band matrix obtained from C by replacing its entries outside the p-band by zeros.

The Modified Block Incomplete Decomposition (MBID) Algorithm

Set $\hat{U}_1 = A_{11}$. Compute for $j = 2, 3, \ldots, m$

$$\begin{aligned} U_j &= A_{jj} - A_{j,j-1}\,\mathrm{inv}(U_{j-1})\,A_{j-1,j}, \\ \tilde{U}_j &= U_j^{(p)}, \\ \mathbf{r}_j &= \mathrm{rs}\big(A_{jj} - A_{j,j-1}\hat{U}_{j-1}^{-1}A_{j-1,j} - \tilde{U}_j\big). \\ \hat{U}_j &= \tilde{U}_j + \mathrm{diag}[\mathbf{r}_j]. \end{aligned} \tag{2}$$

The preconditioner for (1) is then found as follows: Set $\mathbf{y}_1 = \mathbf{b}_1$. For $j = 1, 2, \ldots, m$ compute

$$\mathbf{y}_j = \mathbf{b}_j - A_{j,j-1}\hat{U}_{j-1}^{-1}\mathbf{y}_{j-1}.$$

Set $\mathbf{x}_m = \hat{U}_m^{-1}\mathbf{y}_m$. Compute $j = m-1, m-2, \ldots, 1$

$$\mathbf{x}_j = \hat{U}_j^{-1}(\mathbf{y}_j - A_{j,1+j}\mathbf{y}_{j+1}).$$

Thus, the algorithm relies on three ideas: replacing U_j by its approximate inverse, "cutting" the diagonals outside the chosen p-band, and the row-sum compensation of the error incurred. In the simplest case, we use the approximation

$$\operatorname{inv}(U_j) = (2I - D_j^{-1}U_j)D_j^{-1}, \tag{3}$$

where D_j denotes the diagonal part of U_j. In particular, this can be done when the matrix A satisfies the following conditions:

(H1) The off-diagonal elements of A are nonpositive.

(H2) A is weakly diagonally dominant.

(H3) Each row of $A_{j,j+1}$, $1 \leqslant j \leqslant m-1$, has at least one nonzero element.

Obviously, A satisfying (H1)–(H3) is a nonsingular M-matrix.

THEOREM 1. *Let A obey* (H1)–(H3). *Then the matrices $\hat{U}_j$ in* (2) *are strictly diagonally dominant M-matrices.*

Notice that Theorem 1 for a similar algorithm applied to a symmetric matrix can be found in [2]. It asserts, in fact, that the algorithm does not break down and that the matrix in (3) gives a good approximation of the inverse when A satisfies the hypotheses (H1)–(3).

For the symmetric case, we have

THEOREM 2. *Let A obey* (H1)–(H2) *and be symmetric. Then the actual matrix $\tilde{A}$ produced by the MBID algorithm is positive definite. Furthermore, the conditioned matrix $\tilde{A}^{-1}A$ possesses only real positive eigenvalues greater than or equal to* 1.

Proceeding to estimating the quantity

$$K(\tilde{A}^{-1}A) = \frac{\lambda_{\max}(\tilde{A}^{-1}A)}{\lambda_{\min}(\tilde{A}^{-1}A)}, \tag{4}$$

we first note that it can be viewed as the spectral condition number of $\tilde{A}^{-1}A$ in regard to the norm $\|\mathbf{x}\|_A = \mathbf{x}^T A\mathbf{x}$. Confining our attention to the matrices of the form under consideration with $A_{ij} = -\frac{1}{4}I$, $|i-j| = 1$, we distinguish two cases:

$$A_{jj} = T := \begin{bmatrix} 1 & -\frac{1}{4} & & \\ -\frac{1}{4} & 1 & \ddots & \\ & \ddots & \ddots & -\frac{1}{4} \\ & & -\frac{1}{4} & 1 \end{bmatrix}, \qquad j = 1,2,\ldots,m,$$

which produces the matrix A_T corresponding to the Dirichlet problem on a rectangle, and the matrix A_R associated with the Neumann problem and defined by

$$A_{11} = A_{mm} = R + \tfrac{1}{4}I, \qquad A_{jj} = R + \tfrac{1}{2}I, \quad j = 2,3,\ldots,m-1,$$

where

$$R = T - \operatorname{diag}\left[\tfrac{3}{4}, \tfrac{1}{2}, \ldots, \tfrac{1}{2}, \tfrac{3}{4}\right].$$

THEOREM. *Let $\tilde{A}_T$ be produced by the MBID algorithm with $p = 3$ and no step-by-step scaling applied to A_T. Then*

$$K(\tilde{A}_T^{-1}A_T) \leqslant m.$$

THEOREM 4. *If $\tilde{A}_R$ is generated by the MBID algorithm with $p = 3$ and no step-by-step scaling is applied to $A_R + (\alpha/m)S$, $S = diag[I,0,\ldots,0]$, $\alpha > 0$, then*

$$K\left(\tilde{A}_R^{-1}\left(A_R + \frac{\alpha}{m}I\right)\right) \leqslant \frac{m}{r}, \qquad r = \frac{8\alpha}{15(1+4\alpha)}.$$

The latter results show that the condition number in (4) for both problems after preconditioning becomes $O(h^{-1})$ instead of $O(h^{-2})$, where h stands for the mesh spacing. Moreover, the block method suggested turns out

to be superior to the pointwise incomplete decomposition for rectangular problems where $m < n$. In this case, we have the estimate $O(m)$ for the condition number instead of $O(n)$ obtained in [4].

Proofs and a detailed discussion of the results presented here can be found in [6].

REFERENCES

1 O. Axelsson, A survey of preconditioned iterative methods for linear systems of algebraic equations, *BIT* 24:166–187 (1985).
2 P. Concus, G. H. Golub, and G. Meurant, Block preconditioning for the conjugate gradient methods, *SIAM J. Sci. Statist. Comput.*, 6, *No.* 1 (1985).
3 I. Efrat and M. Tismenetsky, Parallel iterative linear solvers for oil reservoir models, *IBM J. Res. and Develop.*, (in press).
4 I. Gustafsson, A class of first order factorization methods, *BIT* 18:142–156 (1978).
5 O. G. Johnson, C. A. Miccelli, and G. Paul, Polynomial preconditioning for conjugate gradient calculations, *SIAM J. Numer. Anal.* 20:362–376 (1983).
6 M. Tismenetsky and I. Efrat, An efficient preconditioning algorithm and its analysis, *IBM Haifa Sc. Tech.* Report #172, November 1985.

SOME THEOREMS IN MATRIX THEORY USING OPTIMIZATION

by HENRY WOLKOWICZ[36]

Many classical inequalities and theorems on matrices can be derived using optimization techniques. First, one formulates the inequality, or the matrix result, as the maximum (or minimum) of a function subject to appropriate constraints. The solution of the optimization problem then provides a proof of the desired result. For example, the Rayleigh principle can be proved by applying the classical Euler-Lagrange multiplier rule of calculus to the optimization problem

$$\text{maximize } \{(x, Ax) : \|x\| = 1,\ x \in R^n\},$$

where A is an n-by-n Hermitian matrix and $(\ ,\ )$ denotes inner product. One obtains the optimal value $\lambda_{\max}$, the largest eigenvalue of A, which is attained by the corresponding eigenvector x. Similarly, one can prove the arithmetic–geometric-mean inequality by applying the Karush-Kuhn-Tucker conditions (e.g. [27]) to the problem

$$\text{maximize } \left\{ \alpha_1 \cdots \alpha_n : \alpha_i \geqslant 0,\ \sum \alpha_i = 1 \right\}.$$

[36] Department of Mathematical Sciences, University of Delaware, Newark, DE 19716.

Other inequalities, such as Holder's and the Kantorovitch inequality, can be proved in this way; see e.g. [22,33]. Matrix results such as the Perron-Frobenius theorem and the fact that the inverse of an M-matrix is nonnegative can also be proved using optimization; see e.g. [4,32]. In [10], we characterize the Hermitian positive definite matrix which maximizes the determinant given that certain of the elements are prespecified.

The above techniques can also be used to generate new inequalities and matrix results. For example, suppose that A is Hermitian with eigenvalues

$$\lambda_1 \geqslant \cdots \geqslant \lambda_n .$$

Let

$$m := \frac{\operatorname{tr} A}{n}, \qquad s^2 := \frac{\operatorname{tr} A^2}{n} - m^2.$$

By finding the explicit solution of the optimization problem

$$\text{minimize } \lambda_k$$

subject to

(a) $$\sum_{i=1}^{n} \lambda_i = \operatorname{tr} A,$$

(b) $$\sum_{i=1}^{n} \lambda_i^2 \leqslant \operatorname{tr} A^2,$$

(c) $$\lambda_k - \lambda_i \leqslant 0, \qquad i = 1, \ldots, k-1,$$

(d) $$\lambda_i - \lambda_k \leqslant 0, \qquad i = k+1, \ldots, n,$$

one derives the lower bound

$$\lambda_k \geqslant m - s\left(\frac{k-1}{n-k+1}\right)^{1/2},$$

with equality if and only if

$$\lambda_1 = \cdots = \lambda_{k-1} \quad \text{and} \quad \lambda_k = \cdots = \lambda_n.$$

Once the inequality is known, new and simpler proofs can be found. This

approach is taken in [29], where upper and lower bounds are derived for

$$\sum_{i=k}^{l} \lambda_i, \quad \lambda_k - \lambda_l, \quad \frac{\lambda_k}{\lambda_l}.$$

These bounds are in terms of the first two moments of the eigenvalues, i.e. in terms of $\operatorname{tr} A$ and $\operatorname{tr} A^2$. These bounds are then extended to the case of complex eigenvalues, i.e., corresponding bounds are obtained for functions of the real part, imaginary part, and modulus of the eigenvalues. These bounds are then improved, by using measures of nonnormality, in [30].

The bounds for the real eigenvalues in [29] involve the first two moments $\operatorname{tr} A$ and $\operatorname{tr} A^2$. They can be viewed as bounds for the elements of an ordered sample when the mean m and the variance s^2 are known. In this guise, several of the bounds in [29] have reappeared periodically in the matrix-theory as well as in the statistical literature. In particular, the bound for the largest eigenvalue,

$$\lambda_1 \leqslant m + s(n-1)^{1/2},$$

has appeared several times. (See [25] for a survey. Several references are included below.) The λ_i can also be viewed as the roots of a polynomial. The bounds are then given in terms of the first two symmetric functions or the first two coefficients of the (monic) polynomial.

Deriving the bounds using Lagrange-multiplier techniques can be very hard and tedious. In [5, 11, 16, 18, 33] a perturbation technique is introduced. This perturbation technique can be used to improve many of the bounds by introducing additional constraints, such as the fact that the vector of eigenvalues majorizes the diagonal of a Hermitian matrix; see e.g. [33]. One can also improve the bounds by adding more moments, e.g. $\operatorname{tr} A^3$.

REFERENCES

1 Barry C. Arnold, Schwartz, regression, and extreme deviance, *Amer. Statist.* 28:22–23 (1974).

2 P. R. Beesack, On bounds for the range of ordered variates, *Publ. Fac. Electrotechnique Univ. Belgrade Ser. Math. et Phys.* 428:93–96 (1973).

3 P. R. Beesack, On bounds for the range of ordered variates II, *Aequationes Math.* 14:293–301 (1976).

4 J. M. Borwein and H. Wolkowicz, Matrix theorems using optimization, in preparation.
5 J. M. Borwein, G. P. H. Styan, and H. Wolkowicz, Some inequalities involving statistical expressions, *SIAM Rev.* 24:340–342 (1982).
6 A. V. Boyd, Bounds for order statistics, *Publ. Fac. Electrotechnique Univ. Belgrade, Ser. Math. et Phys.* 365:31–32.
7 H. D. Brunk, Note on two papers of K. R. Nair, *J. Indian Soc. Agricultural Statist.* 11:186–189 (1959).
8 A. Clausing, Kantorvich-type inequalities, *Amer. Math. Monthly* 89:314–330 (1982).
9 Meyer Dwass, The extreme deviations inequality, *Amer. Statist.* 29:108 (1975).
10 R. Grone, C. R. Johnson, E. M. Sa, and H. Wolkowicz, Positive definite completions of partial Hermitian matrices, *Linear Algebra Appl.* 58:109–124 (1984).
11 R. Grone, C. R. Johnson, E. M. Sa, and H. Wolkowicz, Improving Hadamard's inequality, *Linear and Multilinear Algebra* 16:305–322 (1984).
12 Douglas M. Hawkins, On the bounds of the range of order statistics, *J. Amer. Statist. Assoc.* 66:644–645 (1971).
13 D. G. Kabe, On extensions of Samuelson's inequality, *Amer. Statist.* 34:249 (1980).
14 C. L. Mallows and Donald Richter, Inequalities of Chebyshev type involving conditional expectations, *Ann. Math. Statist.* 40:1922–1932 (1969).
15 A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*, Academic, 1979.
16 J. Merikoski, G. P. H. Styan, and H. Wolkowicz, Bounds for ratios of eigenvalues using traces, *Linear Algebra Appl* 55:105–124 (1983).
17 J. K. Merikoski and H. Wolkowicz, Improving eigenvalue bounds using extra bounds, *Linear Algebra Appl* 68:93–113 (1985).
18 K. R. Nair, Certain symmetrical properties of unbiased estimates of variance and covariance, *J. Indian Soc. Agricultural Statist.* 1:162–172 (1948).
19 K. R. Nair, A tail-piece to Brunk's paper, *J. Indian Soc. Agricultural Statist.* 11:189–190 (1959).
20 Jagdish Patel, C. H. Kapadia, and D. B. Owen, *Handbook of Statistical Distributions*, Marcel Dekker, New York, 1976.
21 E. S. Pearson and Chandra C. Sekar, The efficiency of statistical tools and a criterion for the rejection of outlying observations, *Biometrika*, 27:308–320 (1936).
22 B. H. Pourciau, Modern multiplier rules, *Amer. Math. Monthly* 87:443–452 (1980).
23 C. P. Quesenberry, The extreme deviations inequality, *Amer. Statist.* 28:112 (1974).
24 Paul A. Samuelson, How deviant can you be?, *J. Amer. Statist. Assoc.* 63:1522–1525 (1968).
25 G. P. H. Styan and H. Wolkowicz, Samuelson's inequality, in *Encyclopedia of Statistical Sciences*, to appear.

26 W. R. Thompson, On a criterion for the rejection of observations and the distribution of the ratio of deviation to sample standard deviation, *Ann. Math. Statist.* 6:214–219 (1935).
27 H. W. Tucker, Nonlinear programming: A historical view, *SIAM-AMS Proc.* 9:1–26 (1976).
28 H. Wolkowicz and G. P. H. Styan, Extensions of Samuelson's inequality, *Amer. Statist* 33:143–144 (1979).
29 H. Wolkowicz and G. P. H. Styan, Bounds for eigenvalues using traces, *Linear Algebra Appl.* 29:471–506 (1980).
30 H. Wolkowicz and G. P. H. Styan, More bounds for eigenvalues using traces, *Linear Algebra Appl.* 31:1–18 (1980).
31 H. Wolkowicz and G. P. H. Styan, Reply, Letter Section, *Amer. Statist.* 34:250 (1984).
32 H. Wolkowicz, Some applications of optimization in matrix theory, *Linear Algebra Appl.* 40:101–118 (1981).
33 H. Wolkowicz, Generating eigenvalue bounds using optimization, Research Report, Univ. of Maryland, College Park, Md. 1982.

Received 27 January 1986