

Fast Reliable Algorithms for Matrices with Structure

Fast Reliable Algorithms for Matrices with Structure

Edited by

T. Kailath
Stanford University
Stanford, California

A. H. Sayed
University of California
Los Angeles, California



Society for Industrial and Applied Mathematics
Philadelphia

Copyright © 1999 by Society for Industrial and Applied Mathematics.

10 9 8 7 6 5 4 3 2 1

All rights reserved. Printed in the United States of America. No part of this book may be reproduced, stored, or transmitted in any manner without the written permission of the publisher. For information, write to the Society for Industrial and Applied Mathematics, 3600 University City Science Center, Philadelphia, PA 19104-2688.

Library of Congress Cataloging-in-Publication Data

Fast reliable algorithms for matrices with structure / edited by
T. Kailath, A.H. Sayed

p. cm.

Includes bibliographical references (p. -) and index.

ISBN 0-89871-431-1 (pbk.)

1. Matrices -- Data processing. 2. Algorithms. I. Kailath,
Thomas. II. Sayed, Ali H.

QA188 .F38 1999

512.9'434 -- dc21

99-26368

CIP

rev.

siam is a registered trademark.

CONTRIBUTORS

Dario A. BINI
Dipartimento di Matematica
Università di Pisa
Pisa, Italy

Sheryl BRANHAM
Dept. Math. and Computer Science
Lehman College
City University of New York
New York, NY 10468, USA

Richard P. BRENT
Oxford University Computing Laboratory
Wolfson Building, Parks Road
Oxford OX1 3QD, England

Raymond H. CHAN
Department of Mathematics
The Chinese University of Hong Kong
Shatin, Hong Kong

Shivkumar CHANDRASEKARAN
Dept. Electrical and Computer Engineering
University of California
Santa Barbara, CA 93106, USA

Patrick DEWILDE
DIMEs, POB 5031, 2600GA Delft
Delft University of Technology
Delft, The Netherlands

Victor S. GRIGORASCU
Facultatea de Electronica and Telecomunicatii
Universitatea Politehnica Bucuresti
Bucharest, Romania

Thomas KAILATH
Department of Electrical Engineering
Stanford University
Stanford, CA 94305, USA

Beatrice MEINI
Dipartimento di Matematica
Università di Pisa
Pisa, Italy

Victor Y. PAN
Dept. Math. and Computer Science
Lehman College
City University of New York
New York, NY 10468, USA

Michael K. NG
Department of Mathematics
The University of Hong Kong
Pokfulam Road, Hong Kong

Phillip A. REGALIA
Signal and Image Processing Dept.
Inst. National des Télécommunications
F-91011 Evry cedex, France

Rhys E. ROSHOLT
Dept. Math. and Computer Science
City University of New York
Lehman College
New York, NY 10468, USA

Ali H. SAYED
Electrical Engineering Department
University of California
Los Angeles, CA 90024, USA

Paolo TILLI
Scuola Normale Superiore
Piazza Cavalieri 7
56100 Pisa, Italy

Ai-Long ZHENG
Department of Mathematics
City University of New York
New York, NY 10468, USA

CONTENTS

PREFACE	xiii
NOTATION	xv
1 DISPLACEMENT STRUCTURE AND ARRAY ALGORITHMS	1
<i>Thomas Kailath</i>	
1.1 Introduction	1
1.2 Toeplitz Matrices	2
1.3 Versions of Displacement Structure	6
1.4 Application to the Fast Evaluation of Matrix-Vector Products	12
1.5 Two Fundamental Properties	13
1.6 Fast Factorization of Structured Matrices	15
1.6.1 Schur Reduction (Deflation)	15
1.6.2 Close Relation to Gaussian Elimination	16
1.6.3 A Generalized Schur Algorithm	17
1.6.4 Array Derivation of the Algorithm	18
1.6.5 An Elementary Section	23
1.6.6 A Simple Example	24
1.7 Proper Form of the Fast Algorithm	25
1.7.1 Positive Lengths	25
1.7.2 Negative Lengths	26
1.7.3 Statement of the Algorithm in Proper Form	26
1.7.4 An Associated Transmission Line	27
1.7.5 Shift-Structured ($F = Z$) Positive-Definite Matrices	27
1.7.6 The Classical Schur Algorithm	28
1.7.7 The Special Case of Toeplitz Matrices	30
1.8 Some Applications in Matrix Computation	30
1.8.1 Going Beyond $F = Z$	31
1.8.2 Simultaneous Factorization of T and T^{-1}	31
1.8.3 QR Factorization of Structured Matrices	32
1.8.4 Avoiding Back Substitution in Linear Equations	33
1.9 Look-Ahead (Block) Schur Algorithm	34
1.10 Fast Inversion of Structured Matrices	37
1.10.1 Schur Construction (Inflation)	37
1.10.2 Statement of the Fast Inversion Algorithm	39
1.10.3 Algorithm Development	40
1.10.4 An Example	41

1.10.5	Proper Form of the Algorithm	42
1.10.6	Application to Pick Matrices	44
1.10.7	The Degenerate Case	44
1.11	Non-Hermitian Toeplitz-Like Matrices	45
1.12	Schur Algorithm for Hankel-Like Matrices	47
1.13	Structured Matrices and Pivoting	48
1.13.1	Incorporating Pivoting into Generalized Schur Algorithms	49
1.13.2	Transformations to Cauchy-Like Structures	50
1.13.3	Numerical Issues	51
1.14	Some Further Issues	52
1.14.1	Incorporating State-Space Structure	52
1.14.2	Iterative Methods	53
1.14.3	Interpolation Problems	54
1.14.4	Inverse Scattering	55
2	STABILIZED SCHUR ALGORITHMS	57
	<i>Shivkumar Chandrasekaran and Ali H. Sayed</i>	
2.1	Introduction	57
2.2	Contributions and Problems	58
2.3	Related Works in the Literature	58
2.3.1	Notation	60
2.3.2	Brief Review of Displacement Structure	60
2.4	The Generalized Schur Algorithm	62
2.5	A Limit to Numerical Accuracy	64
2.6	Implementations of the Hyperbolic Rotation	65
2.6.1	Direct Implementation	66
2.6.2	Mixed Downdating	66
2.6.3	The Orthogonal-Diagonal Procedure	67
2.6.4	The H Procedure	68
2.7	Implementation of the Blaschke Matrix	69
2.7.1	The Case of Diagonal F	69
2.7.2	The Case of Strictly Lower Triangular F	70
2.8	Enforcing Positive Definiteness	71
2.9	Accuracy of the Factorization	71
2.10	Pivoting with Diagonal F	73
2.10.1	A Numerical Example	74
2.10.2	The Case of Positive F	74
2.10.3	The Nonpositive Case	74
2.11	Controlling the Generator Growth	75
2.12	Solution of Linear Systems of Equations	75
2.12.1	Computing the Matrix R	75
2.12.2	Iterative Refinement	76
2.13	Enhancing the Robustness of the Algorithm	77
2.13.1	Enhancing the OD Method	77
2.13.2	Enhancing the H Procedure	78
2.13.3	A Numerical Example	78
2.14	Results for the Positive-Definite Case	79
2.A	Pseudocode for the Stable Schur Algorithm	81

3	FAST STABLE SOLVERS FOR STRUCTURED LINEAR SYSTEMS	85
	<i>Ali H. Sayed and Shivkumar Chandrasekaran</i>	
3.1	Introduction	85
3.2	Overview of the Proposed Solution	86
3.3	The Generalized Schur Algorithm for Indefinite Matrices	87
3.4	Fast QR Factorization of Shift-Structured Matrices	90
3.4.1	The Toeplitz Case	91
3.4.2	Other Augmentations	92
3.5	Well-Conditioned Coefficient Matrices	93
3.5.1	Implementation of the Hyperbolic Rotation	94
3.5.2	Avoiding Breakdown	95
3.5.3	Error Analysis of the Last Steps	96
3.5.4	Summary	97
3.5.5	Solving the Linear System of Equations	97
3.6	Ill-Conditioned Coefficient Matrices	97
3.6.1	Solving the Linear System of Equations	98
3.6.2	Conditions on the Coefficient Matrix	99
3.7	Summary of the Algorithm	99
4	STABILITY OF FAST ALGORITHMS FOR STRUCTURED LINEAR SYSTEMS	103
	<i>Richard P. Brent</i>	
4.1	Introduction	103
4.1.1	Outline	104
4.1.2	Notation	105
4.2	Stability and Weak Stability	105
4.2.1	Gaussian Elimination with Pivoting	106
4.2.2	Weak Stability	107
4.2.3	Example: Orthogonal Factorization	108
4.3	Classes of Structured Matrices	108
4.3.1	Cauchy and Cauchy-Like Matrices	109
4.3.2	Toeplitz Matrices	109
4.4	Structured Gaussian Elimination	109
4.4.1	The GKO-Toeplitz Algorithm	110
4.4.2	Error Analysis	111
4.4.3	A General Strategy	112
4.5	Positive-Definite Structured Matrices	112
4.5.1	The Bareiss Algorithm for Positive-Definite Matrices	113
4.5.2	Generalized Schur Algorithms	113
4.6	Fast Orthogonal Factorization	114
4.6.1	Use of the Seminormal Equations	115
4.6.2	Computing Q Stably	115
4.6.3	Solution of Indefinite or Unsymmetric Structured Systems	115
4.7	Concluding Remarks	116
5	ITERATIVE METHODS FOR LINEAR SYSTEMS WITH MATRIX STRUCTURE	117
	<i>Raymond H. Chan and Michael K. Ng</i>	
5.1	Introduction	117
5.2	The CG Method	118

5.3	Iterative Methods for Solving Toeplitz Systems	121
5.3.1	Preconditioning	122
5.3.2	Circulant Matrices	123
5.3.3	Toeplitz Matrix-Vector Multiplication	124
5.3.4	Circulant Preconditioners	125
5.4	Band-Toeplitz Preconditioners	130
5.5	Toeplitz-Circulant Preconditioners	132
5.6	Preconditioners for Structured Linear Systems	133
5.6.1	Toeplitz-Like Systems	133
5.6.2	Toeplitz-Plus-Hankel Systems	137
5.7	Toeplitz-Plus-Band Systems	139
5.8	Applications	140
5.8.1	Linear-Phase Filtering	140
5.8.2	Numerical Solutions of Biharmonic Equations	142
5.8.3	Queueing Networks with Batch Arrivals	144
5.8.4	Image Restorations	147
5.9	Concluding Remarks	149
5.A	Proof of Theorem 5.3.4	150
5.B	Proof of Theorem 5.6.2	151
6	ASYMPTOTIC SPECTRAL DISTRIBUTION OF TOEPLITZ-RELATED MATRICES	153
	<i>Paolo Tilli</i>	
6.1	Introduction	153
6.2	What Is Spectral Distribution?	153
6.3	Toeplitz Matrices and Shift Invariance	157
6.3.1	Spectral Distribution of Toeplitz Matrices	158
6.3.2	Unbounded Generating Function	162
6.3.3	Eigenvalues in the Non-Hermitian Case	163
6.3.4	The Szegő Formula for Singular Values	164
6.4	Multilevel Toeplitz Matrices	166
6.5	Block Toeplitz Matrices	170
6.6	Combining Block and Multilevel Structure	174
6.7	Locally Toeplitz Matrices	175
6.7.1	A Closer Look at Locally Toeplitz Matrices	178
6.7.2	Spectral Distribution of Locally Toeplitz Sequences	182
6.8	Concluding Remarks	186
7	NEWTON'S ITERATION FOR STRUCTURED MATRICES	189
	<i>Victor Y. Pan, Sheryl Branham, Rhys E. Rosholt, and Ai-Long Zheng</i>	
7.1	Introduction	189
7.2	Newton's Iteration for Matrix Inversion	190
7.3	Some Basic Results on Toeplitz-Like Matrices	192
7.4	The Newton-Toeplitz Iteration	194
7.4.1	Bounding the Displacement Rank	195
7.4.2	Convergence Rate and Computational Complexity	196
7.4.3	An Approach Using f -Circulant Matrices	198
7.5	Residual Correction Method	200
7.5.1	Application to Matrix Inversion	200
7.5.2	Application to a Linear System of Equations	201

7.5.3	Application to a Toeplitz Linear System of Equations	201
7.5.4	Estimates for the Convergence Rate	203
7.6	Numerical Experiments	204
7.7	Concluding Remarks	207
7.A	Correctness of Algorithm 7.4.2	208
7.B	Correctness of Algorithm 7.5.1	209
7.C	Correctness of Algorithm 7.5.2	209
8	FAST ALGORITHMS WITH APPLICATIONS TO MARKOV CHAINS AND QUEUEING MODELS	211
	<i>Dario A. Bini and Beatrice Meini</i>	
8.1	Introduction	211
8.2	Toeplitz Matrices and Markov Chains	212
8.2.1	Modeling of Switches and Network Traffic Control	214
8.2.2	Conditions for Positive Recurrence	215
8.2.3	Computation of the Probability Invariant Vector	216
8.3	Exploitation of Structure and Computational Tools	217
8.3.1	Block Toeplitz Matrices and Block Vector Product	218
8.3.2	Inversion of Block Triangular Block Toeplitz Matrices	221
8.3.3	Power Series Arithmetic	223
8.4	Displacement Structure	224
8.5	Fast Algorithms	226
8.5.1	The Fast Ramaswami Formula	227
8.5.2	A Doubling Algorithm	227
8.5.3	Cyclic Reduction	230
8.5.4	Cyclic Reduction for Infinite Systems	234
8.5.5	Cyclic Reduction for Generalized Hessenberg Systems	239
8.6	Numerical Experiments	241
9	TENSOR DISPLACEMENT STRUCTURES AND POLYSPECTRAL MATCHING	245
	<i>Victor S. Grigorascu and Phillip A. Regalia</i>	
9.1	Introduction	245
9.2	Motivation for Higher-Order Cumulants	245
9.3	Second-Order Displacement Structure	249
9.4	Tucker Product and Cumulant Tensors	251
9.5	Examples of Cumulants and Tensors	254
9.6	Displacement Structure for Tensors	257
9.6.1	Relation to the Polyspectrum	258
9.6.2	The Linear Case	261
9.7	Polyspectral Interpolation	264
9.8	A Schur-Type Algorithm for Tensors	268
9.8.1	Review of the Second-Order Case	268
9.8.2	A Tensor Outer Product	269
9.8.3	Displacement Generators	272
9.9	Concluding Remarks	275

10 MINIMAL COMPLEXITY REALIZATION OF STRUCTURED MATRICES	277
<i>Patrick Dewilde</i>	
10.1 Introduction	277
10.2 Motivation of Minimal Complexity Representations	278
10.3 Displacement Structure	279
10.4 Realization Theory for Matrices	280
10.4.1 Nerode Equivalence and Natural State Spaces	283
10.4.2 Algorithm for Finding a Realization	283
10.5 Realization of Low Displacement Rank Matrices	286
10.6 A Realization for the Cholesky Factor	289
10.7 Discussion	293
A USEFUL MATRIX RESULTS	297
<i>Thomas Kailath and Ali H. Sayed</i>	
A.1 Some Matrix Identities	298
A.2 The Gram–Schmidt Procedure and the QR Decomposition	303
A.3 Matrix Norms	304
A.4 Unitary and \mathcal{J} -Unitary Transformations	305
A.5 Two Additional Results	306
B ELEMENTARY TRANSFORMATIONS	309
<i>Thomas Kailath and Ali H. Sayed</i>	
B.1 Elementary Householder Transformations	310
B.2 Elementary Circular or Givens Rotations	312
B.3 Hyperbolic Transformations	314
BIBLIOGRAPHY	321
INDEX	339

PREFACE

The design of fast and numerically reliable algorithms for large-scale matrix problems with structure has become an increasingly important activity, especially in recent years, driven by the ever-increasing complexity of applications arising in control, communications, computation, and signal processing.

The major challenge in this area is to develop algorithms that blend speed and numerical accuracy. These two requirements often have been regarded as competitive, so much so that the design of fast and numerically reliable algorithms for large-scale structured linear matrix equations has remained a significant open issue in many instances.

This problem, however, has been receiving increasing attention recently, as witnessed by a series of international meetings held in the last three years in Santa Barbara (USA, Aug. 1996), Cortona (Italy, Sept. 1996), and St. Emilion (France, Aug. 1997). These meetings provided a forum for the exchange of ideas on current developments, trends, and issues in fast and reliable computing among peer research groups. The idea of this book project grew out of these meetings, and the chapters are selections from works presented at the meetings. In the process, several difficult decisions had to be made; the editors beg the indulgence of participants whose contributions could not be included here.

Browsing through the chapters, the reader soon will realize that this project is unlike most edited volumes. The book is not merely a collection of submitted articles; considerable effort went into blending the several chapters into a reasonably consistent presentation. We asked each author to provide a contribution with a significant tutorial value. In this way, the chapters not only provide the reader with an opportunity to review some of the most recent advances in a particular area of research, but they do so with enough background material to put the work into proper context. Next, we carefully revised and revised again each submission to try to improve both clarity and uniformity of presentation. This was a substantial undertaking since we often needed to change symbols across chapters, to add cross-references to other chapters and sections, to reorganize sections, to reduce redundancy, and to try to state theorems, lemmas, and algorithms uniformly across the chapters. We did our best to ensure a uniformity of presentation and notation but, of course, errors and omissions may still exist and we apologize in advance for any of these. We also take this opportunity to thank the authors for their patience and for their collaboration during this time-consuming process. In all we believe the book includes a valuable collection of chapters that cover in some detail different aspects of the most recent trends in the theory of fast algorithms, with emphasis on implementation and application issues.

The book may be divided into four distinct parts:

1. The first four chapters deal with fast *direct* methods for the triangular factorization

of structured matrices, as well as the solution of structured linear systems of equations. The emphasis here is mostly on the generalized Schur algorithm, its numerical properties, and modifications to ensure numerical stability.

2. Chapters 5, 6, and 7 deal with fast *iterative* methods for the solution of structured linear systems of equations. The emphasis here is on the preconditioned conjugate gradient method and on Newton's method.
3. Chapters 8 to 10 deal with extensions of the notion of structure to the block case, the tensor case, and to the input-output framework. Chapter 8 presents fast algorithms for block Toeplitz systems of equations and considers applications in Markov chains and queueing theory. Chapter 9 studies tensor displacement structure and applications in polyspectral interpolation. Chapter 10 discusses realization theory and computational models for structured problems.
4. We have included two appendices that collect several useful matrix results that are used in several places in the book.

Acknowledgments. We gratefully acknowledge the support of the Army Research Office and the National Science Foundation in funding the organization of the Santa Barbara Workshop. Other grants from these agencies, as well as from the Defense Advanced Research Projects Agency and the Air Force Office of Scientific Research, supported the efforts of the editors on this project. We are also grateful to Professors Alan Laub of University of California Davis and Shivkumar Chandrasekaran of University of California Santa Barbara for their support and joint organization with the editors of the 1996 Santa Barbara Workshop. It is also a pleasure to thank Professors M. Najim of the University of Bordeaux and P. Dewilde of Delft University, for their leading role in the St. Emilion Workshop, and Professor D. Bini of the University of Pisa and several of his Italian colleagues, for the fine 1996 Toeplitz Workshop in Cortona.

October 1998

T. Kailath
Stanford, CA

A. H. Sayed
Westwood, CA

NOTATION

\mathbf{N}	The set of natural numbers.
\mathbf{Z}	The set of integer numbers.
\mathbf{R}	The set of real numbers.
\mathbf{C}	The set of complex numbers.
\emptyset	The empty set.
$\mathcal{C}_{2\pi}$	The set of 2π -periodic complex-valued continuous functions defined on $[-\pi, \pi]$.
$\mathcal{C}_0(\mathbf{R})$	The set of complex-valued continuous functions with bounded support in \mathbf{R} .
$\mathcal{C}_b(\mathbf{R})$	The set of bounded and uniformly continuous complex-valued functions over \mathbf{R} .
\cdot^T	Matrix transposition.
\cdot^*	Complex conjugation for scalars and conjugate transposition for matrices.
$a \triangleq b$	The quantity a is defined as b .
$\text{col}\{a, b\}$	A column vector with entries a and b .
$\text{diag}\{a, b\}$	A diagonal matrix with diagonal entries a and b .
$\text{tridiag}\{a, b, c\}$	A tridiagonal Toeplitz matrix with b along its diagonal, a along its lower diagonal, and c along its upper diagonal.
$a \oplus b$	The same as $\text{diag}\{a, b\}$.
$\hat{\mathbf{i}}$	$\sqrt{-1}$
$\lceil x \rceil$	The smallest integer $m \geq x$.
$\lfloor x \rfloor$	The largest integer $m \leq x$.
$\mathbf{0}$	A zero scalar, vector, or matrix.
\mathbf{I}_n	The identify matrix of size $n \times n$.
$\mathcal{L}(x)$	A lower triangular Toeplitz matrix whose first column is x .
\diamond	The end of a proof, an example, or a remark.

$\ \cdot\ _2$	The Euclidean norm of a vector or the maximum singular value of a matrix.
$\ \cdot\ _1$	The sum of the absolute values of the entries of a vector or the maximum absolute column sum of a matrix.
$\ \cdot\ _\infty$	The largest absolute entry of a vector or the maximum absolute row sum of a matrix.
$\ \cdot\ _F$	The Frobenius norm of a matrix.
$\ \cdot\ $	Some vector or matrix norm.
$ A $	A matrix with elements $ a_{ij} $.
$\lambda_i(A)$	i th eigenvalue of A .
$\sigma_i(A)$	i th singular value of A .
$\kappa(A)$	Condition number of a matrix A , given by $\ A\ _2\ A^{-1}\ _2$.
$\text{cond}_k(A)$	Equal to $\ A\ _k\ A^{-1}\ _k$.
ε	Machine precision.
$O(n)$	A constant multiple of n , or of the order of n .
$O_n(\varepsilon)$	$O(\varepsilon c(n))$, where $c(n)$ is some polynomial in n .
$\hat{\cdot}$	A computed quantity in a finite precision algorithm.
$\bar{\cdot}$	An intermediate exact quantity in a finite precision algorithm.
CG	The conjugate gradient method.
LDU	The lower-diagonal-upper triangular factorization of a matrix.
PCG	The preconditioned conjugate gradient method.
QR	The QR factorization of a matrix.