

NUMERICAL
METHODS
FOR
LEAST
SQUARES
PROBLEMS

NUMERICAL
METHODS
FOR
LEAST
SQUARES
PROBLEMS

ÅKE BJÖRCK

**Linköping University
Linköping, Sweden**

siam.

Society for Industrial and Applied Mathematics
Philadelphia

Copyright © 1996 by the Society for Industrial and Applied Mathematics

1098765

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 Market Street, 6th Floor, Philadelphia, PA 19104-2688 USA.

Library of Congress Cataloging-in-Publication Data

Björck, Åke, 1934-

Numerical methods for least squares problems / Åke Björck.

p.cm.

Includes bibliographical references (p. -) and index.

ISBN 978-0-898713-60-2 (pbk.)

1. Equations, Simultaneous—Numerical solutions. 2. Least squares. I. Title.

QA214.B56 1996

512.9'42—dc20

96-3908

Portions were adapted with permission from *Handbook of Numerical Analysis, Volume I, Least Squares Methods* by Åke Björck, c 1990, North-Holland, Amsterdam.

Dedicated to
Germund Dahlquist
and
Gene H. Golub

Contents

Preface xv

1. Mathematical and Statistical Properties of Least Squares Solutions 1

1.1	Introduction	1
1.1.1	Historical remarks.	2
1.1.2	Statistical preliminaries.	2
1.1.3	Linear models and the Gauss–Markoff theorem.	3
1.1.4	Characterization of least squares solutions.	5
1.2	The Singular Value Decomposition	9
1.2.1	The singular value decomposition.	9
1.2.2	Related eigenvalue decompositions.	11
1.2.3	Matrix approximations.	12
1.2.4	The sensitivity of singular values and vectors.	13
1.2.5	The SVD and pseudoinverse.	15
1.2.6	Orthogonal projectors and angles between subspaces.	17
1.3	The QR Decomposition	19
1.3.1	The full rank case.	19
1.3.2	Rank revealing QR decompositions.	21
1.3.3	The complete orthogonal decomposition.	23
1.4	Sensitivity of Least Squares Solutions	24
1.4.1	Vector and matrix norms.	24
1.4.2	Perturbation analysis of pseudoinverses.	26
1.4.3	Perturbation analysis of least squares solutions.	27
1.4.4	Asymptotic forms and derivatives.	32
1.4.5	Componentwise perturbation analysis.	32
1.4.6	A posteriori estimation of errors.	34

2. Basic Numerical Methods 37

2.1	Basics of Floating Point Computation	37
2.1.1	Rounding error analysis.	37
2.1.2	Running rounding error analysis.	39

2.1.3	Stability of algorithms.	40
2.2	The Method of Normal Equations	42
2.2.1	Forming the normal equations.	42
2.2.2	The Cholesky factorization.	44
2.2.3	Conditioning and scaling.	49
2.3	Elementary Orthogonal Transformations	51
2.3.1	Householder transformations.	51
2.3.2	Givens transformation.	53
2.3.3	Fast Givens transformations.	56
2.4	Methods Based on the QR Decomposition	58
2.4.1	Householder and Givens QR decomposition.	58
2.4.2	Gram–Schmidt orthogonalization.	60
2.4.3	Least squares by Householder QR decomposition.	63
2.4.4	Least squares problems by MGS.	64
2.4.5	Gram–Schmidt with reorthogonalization.	66
2.4.6	Hybrid algorithms.	69
2.4.7	Block algorithms.	71
2.5	Methods Based on Gaussian Elimination	73
2.5.1	The Peters–Wilkinson method.	73
2.5.2	Pseudoinverse solutions from LU decompositions.	76
2.5.3	The augmented system method.	77
2.6	Computing the SVD	81
2.6.1	SVD and least squares problems.	81
2.6.2	Transformation to bidiagonal form.	81
2.6.3	The QR algorithm for real symmetric matrices.	83
2.6.4	The QR algorithm for the SVD.	85
2.6.5	Zero shift QR algorithm.	90
2.6.6	Jacobi methods for the SVD.	92
2.6.7	Singular values by spectrum slicing.	96
2.7	Rank Deficient and Ill-Conditioned Problems	99
2.7.1	SVD and numerical rank.	99
2.7.2	Truncated SVD solutions and regularization.	100
2.7.3	QR decompositions with column pivoting.	103
2.7.4	Pseudoinverse solutions from QR decompositions.	106
2.7.5	Rank revealing QR decompositions.	108
2.7.6	Complete orthogonal decompositions.	110
2.7.7	Subset selection by SVD and RRQR.	113
2.8	Estimating Condition Numbers and Errors	114
2.8.1	The LINPACK condition estimator.	114
2.8.2	Hager’s condition estimator.	116
2.8.3	Computing the variance-covariance matrix.	118
2.9	Iterative Refinement	120
2.9.1	Iterative refinement for linear systems.	120
2.9.2	Extended precision iterative refinement.	121

2.9.3	Fixed precision iterative refinement.	124
3.	Modified Least Squares Problems	127
3.1	Introduction	127
3.1.1	Updating problems.	127
3.1.2	Modified linear systems.	128
3.1.3	Modifying matrix factorizations.	129
3.1.4	Recursive least squares.	131
3.2	Modifying the Full QR Decomposition	132
3.2.1	Introduction.	132
3.2.2	General rank one change.	132
3.2.3	Deleting a column.	133
3.2.4	Appending a column.	135
3.2.5	Appending a row.	136
3.2.6	Deleting a row.	137
3.2.7	Modifying the Gram–Schmidt decomposition.	138
3.3	Downdating the Cholesky Factorization	140
3.3.1	Introduction.	140
3.3.2	The Saunders algorithm.	141
3.3.3	The corrected seminormal equations.	142
3.3.4	Hyperbolic rotations.	143
3.4	Modifying the Singular Value Decomposition	145
3.4.1	Introduction.	145
3.4.2	Appending a row.	145
3.4.3	Deleting a row.	147
3.5	Modifying Rank Revealing QR Decompositions	149
3.5.1	Appending a row.	149
3.5.2	Deleting a row.	152
4.	Generalized Least Squares Problems	153
4.1	Generalized QR Decompositions	153
4.1.1	Introduction.	153
4.1.2	Computing the GQR and PQR.	153
4.2	The Generalized SVD	155
4.2.1	The CS decomposition.	155
4.2.2	The generalized SVD.	157
4.2.3	Computing the GSVD.	159
4.3	General Linear Models and Generalized Least Squares	160
4.3.1	Gauss–Markoff linear models.	160
4.3.2	Generalized linear least squares problems.	162
4.3.3	Paige’s method.	164
4.4	Weighted Least Squares Problems	165
4.4.1	Introduction.	165
4.4.2	Methods based on Gaussian elimination.	166
4.4.3	QR decompositions for weighted problems.	168

4.4.4	Weighted problems by updating.	171
4.5	Minimizing the l_p Norm	172
4.5.1	Introduction.	172
4.5.2	Iteratively reweighted least squares.	173
4.5.3	Robust linear regression.	175
4.5.4	Algorithms for l_1 and l_∞ approximation.	175
4.6	Total Least Squares	176
4.6.1	Errors-in-variables models.	176
4.6.2	Total least squares problem by SVD.	177
4.6.3	Relationship to the least squares solution.	180
4.6.4	Multiple right-hand sides.	181
4.6.5	Generalized TLS problems.	182
4.6.6	Linear orthogonal distance regression.	184
5.	Constrained Least Squares Problems	187
5.1	Linear Equality Constraints	187
5.1.1	Introduction.	187
5.1.2	Method of direct elimination.	188
5.1.3	The nullspace method.	189
5.1.4	Problem LSE by generalized SVD.	191
5.1.5	The method of weighting.	192
5.1.6	Solving LSE problems by updating.	194
5.2	Linear Inequality Constraints	194
5.2.1	Classification of problems.	194
5.2.2	Basic transformations of problem LSI.	196
5.2.3	Active set algorithms for problem LSI.	198
5.2.4	Active set algorithms for BLS.	201
5.3	Quadratic Constraints	203
5.3.1	Ill-posed problems.	203
5.3.2	Quadratic inequality constraints.	205
5.3.3	Problem LSQI by GSVD.	206
5.3.4	Problem LSQI by QR decomposition.	208
5.3.5	Cross-validation.	211
6.	Direct Methods for Sparse Problems	215
6.1	Introduction	215
6.2	Banded Least Squares Problems	217
6.2.1	Storage schemes for banded matrices.	218
6.2.2	Normal equations for banded problems.	219
6.2.3	Givens QR decomposition for banded problems.	221
6.2.4	Householder QR decomposition for banded problems.	222
6.3	Block Angular Least Squares Problems	224
6.3.1	Block angular form.	224
6.3.2	QR methods for block angular problems.	225
6.4	Tools for General Sparse Problems	227

6.4.1	Storage schemes for general sparse matrices.	227
6.4.2	Graph representation of sparse matrices.	230
6.4.3	Predicting the structure of $A^T A$	231
6.4.4	Predicting the structure of R	232
6.4.5	Block triangular form of a sparse matrix.	234
6.5	Fill Minimizing Column Orderings	237
6.5.1	Bandwidth reducing ordering methods.	237
6.5.2	Minimum degree ordering.	238
6.5.3	Nested dissection orderings.	240
6.6	The Numerical Cholesky and QR Decompositions	242
6.6.1	The Cholesky factorization.	242
6.6.2	Row sequential QR decomposition.	242
6.6.3	Row orderings for sparse QR decomposition.	244
6.6.4	Multifrontal QR decomposition.	245
6.6.5	Iterative refinement and seminormal equations.	250
6.7	Special Topics	252
6.7.1	Rank revealing sparse QR decomposition.	252
6.7.2	Updating sparse least squares solutions.	254
6.7.3	Partitioning for out-of-core solution.	255
6.7.4	Computing selected elements of the covariance matrix.	256
6.8	Sparse Constrained Problems	257
6.8.1	An active set method for problem BLS.	257
6.8.2	Interior point methods for problem BLS.	262
6.9	Software and Test Results	264
6.9.1	Software for sparse direct methods.	264
6.9.2	Test results.	266
7.	Iterative Methods For Least Squares Problems	269
7.1	Introduction	269
7.1.1	Iterative versus direct methods.	270
7.1.2	Computing sparse matrix-vector products.	270
7.2	Basic Iterative Methods	274
7.2.1	General stationary iterative methods.	274
7.2.2	Splittings of rectangular matrices.	276
7.2.3	Classical iterative methods.	276
7.2.4	Successive overrelaxation methods.	279
7.2.5	Semi-iterative methods.	280
7.2.6	Preconditioning.	283
7.3	Block Iterative Methods	284
7.3.1	Block column preconditioners.	284
7.3.2	The two-block case.	286
7.4	Conjugate Gradient Methods	288
7.4.1	CGLS and variants.	288
7.4.2	Convergence properties of CGLS.	290

7.4.3	The conjugate gradient method in finite precision.	292
7.4.4	Preconditioned CGLS.	293
7.5	Incomplete Factorization Preconditioners	294
7.5.1	Incomplete Cholesky preconditioners.	294
7.5.2	Incomplete orthogonal decompositions.	297
7.5.3	Preconditioners based on LU factorization.	299
7.6	Methods Based on Lanczos Bidiagonalization	303
7.6.1	Lanczos bidiagonalization.	303
7.6.2	Best approximation in the Krylov subspace.	306
7.6.3	The LSQR algorithm.	307
7.6.4	Convergence of singular values and vectors.	309
7.6.5	Bidiagonalization and total least squares.	310
7.7	Methods for Constrained Problems	312
7.7.1	Problems with upper and lower bounds.	312
7.7.2	Iterative regularization.	314
8.	Least Squares Problems with Special Bases	317
8.1	Least Squares Approximation and Orthogonal Systems	317
8.1.1	General formalism.	317
8.1.2	Statistical aspects of the method of least squares.	318
8.2	Polynomial Approximation	319
8.2.1	Triangle family of polynomials.	319
8.2.2	General theory of orthogonal polynomials.	320
8.2.3	Discrete least squares fitting.	321
8.2.4	Vandermonde-like systems.	323
8.2.5	Chebyshev polynomials.	325
8.3	Discrete Fourier Analysis	328
8.3.1	Introduction.	328
8.3.2	Orthogonality relations.	329
8.3.3	The fast Fourier transform.	330
8.4	Toeplitz Least Squares Problems	332
8.4.1	Introduction.	332
8.4.2	QR decomposition of Toeplitz matrices.	333
8.4.3	Iterative solvers for Toeplitz systems.	334
8.4.4	Preconditioners for Toeplitz systems.	335
8.5	Kronecker Product Problems	336
9.	Nonlinear Least Squares Problems	339
9.1	The Nonlinear Least Squares Problem	339
9.1.1	Introduction.	339
9.1.2	Necessary conditions for local minima.	340
9.1.3	Basic numerical methods.	341
9.2	Gauss-Newton-Type Methods	342
9.2.1	The damped Gauss-Newton method.	343
9.2.2	Local convergence of the Gauss-Newton method.	345

9.2.3	Trust region methods.	346
9.3	Newton-Type Methods	348
9.3.1	Introduction.	348
9.3.2	A hybrid Newton method.	348
9.3.3	Quasi-Newton methods.	349
9.4	Separable and Constrained Problems	351
9.4.1	Separable problems.	351
9.4.2	General constrained problems.	353
9.4.3	Orthogonal distance regression.	354
9.4.4	Least squares fit of geometric elements.	357
Bibliography		359
Index		401

Chapter 1

Preface

A basic problem in science is to fit a model to observations subject to errors. It is clear that the more observations that are available the more accurately will it be possible to calculate the parameters in the model. This gives rise to the problem of “solving” an overdetermined linear or nonlinear system of equations. It can be shown that the solution which minimizes a weighted sum of the squares of the residual is optimal in a certain sense. Gauss claims to have discovered the method of least squares in 1795 when he was 18 years old. Hence this book also marks the bicentennial of the use of the least squares principle.

The development of the basic modern numerical methods for solving linear least squares problems took place in the late sixties. The QR decomposition by Householder transformations was developed by Golub and published in 1965. The implicit QR algorithm for computing the singular value decomposition (SVD) was developed about the same time by Kahan, Golub, and Wilkinson, and the final algorithm was published in 1970. These matrix decompositions have since been developed and generalized to a high level of sophistication. Great progress has been made in the last decade in methods for generalized and modified least squares problems and in direct and iterative methods for large sparse problems. Methods for total least squares problems, which allow errors also in the system matrix, have been systematically developed.

Applications of least squares of crucial importance occur in many areas of applied and engineering research such as statistics, geodetics, photogrammetry, signal processing, and control. Because of the great increase in the capacity for automatic data capturing, least squares problems of large size are now routinely solved. Therefore, sparse direct methods as well as iterative methods play an increasingly important role. Applications in signal processing have created a great demand for stable and efficient methods for modifying least squares solutions when data are added or deleted. This has led to renewed interest in rank revealing QR decompositions, which lend themselves better to updating than the singular value decomposition. Generalized and weighted least squares problems and problems of Toeplitz and Kronecker structure are becoming increasingly important.

Chapter 1 gives the basic facts and the mathematical and statistical background of least squares methods. In Chapter 2 relevant matrix decompositions and basic numerical methods are covered in detail. Although most proofs are omitted, these two chapters are more elementary than the rest of the book and essentially self-contained. Chapter 3 treats modified least squares problems and includes many recent results. In Chapter 4 generalized QR and SVD decompositions are presented, and methods for generalized and weighted problems surveyed. Here also, robust methods and methods for total least squares are treated. Chapter 5 surveys methods for problems with linear and quadratic constraints. Direct and iterative methods for large sparse least squares problems are covered in Chapters 6 and 7. These methods are still subject to intensive research, and the presentation is more advanced. Chapter 8 is devoted to problems with special bases, including least squares fitting of polynomials and problems of Toeplitz and Kronecker structures. Finally, Chapter 9 contains a short survey of methods for nonlinear problems.

This book will be of interest to mathematicians working in numerical linear algebra, computational scientists and engineers, and statisticians, as well as electrical engineers. Although a solid understanding of numerical linear algebra is needed for the more advanced sections, I hope the book will be found useful in upper-level undergraduate and beginning graduate courses in scientific computing and applied sciences.

I have aimed to make the book and the bibliography as comprehensive and up-to-date as possible. Many recent research results are included, which were only available in the research literature before. Inevitably, however, the content reflects my own interests, and I apologize in advance to those whose work has not been mentioned. In particular, work on the least squares problem in the former Soviet Union is, to a large extent, not covered.

The history of this book dates back to at least 1981, when I wrote a survey entitled "Least Squares Methods in Physics and Engineering" for the Academic Training Programme at CERN in Geneva. In 1985 I was invited to contribute a chapter on "Least Squares Methods" in the *Handbook of Numerical Analysis*, edited by P. G. Ciarlet and J. L. Lions. This chapter [95] was finished in 1988 and appeared in Volume 1 of the *Handbook*, published by North-Holland in 1990. The present book is based on this contribution, although it has been extensively updated and made more complete.

The book has greatly benefited from the insight and knowledge kindly provided by many friends and colleagues. In particular, I have been greatly influenced by the work of Gene H. Golub, Nick Higham, and G. W. Stewart. Per-Åke Wedin gave valuable advice on the chapter on nonlinear problems. Part of the *Handbook* chapter was written while I had the benefit of visiting the Division of Mathematics and Statistics at CSIRO in Canberra and the Chr. Michelsen Institute in Bergen.

Thanks are due to Elsevier Science B.V. for the permission to use part of the material from the *Handbook* chapter. Finally, I thank Beth Gallagher and Vickie Kearn at SIAM for the cheerful and professional support they have given throughout the copy editing and production of the book.

Åke Björck
Linköping, February 1996