Unifying Least Squares,
Total Least Squares & Data Least Squares

Chris Paige

School of Computer Science
McGill University
Montreal, Quebec, Canada, H3A 2A7
(chris@cs.mcgill.ca)

and

Zdeněk Strakoš

Institut of Computer Science, Academy of Sciences
Pod vod. vezi 2, 182 07 Praha 8
Czech Republic
(strakos@cs.cas.cz)

Abstract

We usually solve overdetermined linear systems $Bx \approx c$ by constructing minimal corrections to c and/or B such that the corrected system is compatible.

In least squares (LS) the correction is restricted to c, while in data least squares (DLS) it is restricted to B.

Scaled total least squares (STLS) [Rao 1997] (following ideas in [Golub & Van Loan 1980], [Golub Hoffman & Stewart 1987]) allows corrections to both c and B, where the relative size of the corrections depends on a parameter γ.

STLS is a generalization of total least squares (TLS). It becomes TLS when $\gamma = 1$, corresponds to LS when $\gamma \to 0$, and to DLS when $\gamma \to \infty$.

We will discuss here:

- A useful formulation of the STLS problem.
- A new criterion that makes the LS, DLS, TLS & STLS theories consistent.
- The direct reduction of any linear system $Bx \approx c$ to a minimally dimensioned core system satisfying our criterion.
- The basics of practical algorithms for the LS, DLS, TLS and STLS problems.

In Praise of the CS Decomposition

Charlie Van Loan

Computer Science Department
Cornell University
(cv@cs.cornell.edu)

Abstract

Life as we know it would be very different without partitioned orthogonal matrices and what Pete Stewart has to say about them.

Low Rank Circulant Approximation

Bob Plemmons

Department of Computer Science
Wake Forest University
(plemmons@cs.wfu.edu)

(Joint with Moody Chu, NC State University, and interaction with Xiaobai Sun, Duke)

Abstract

Circulant matrices arise in many important applications. The empirical data collected in practice, due to the finite bandwidth of information gathering devices, often do not maintain the circulant structure or induce a certain desired rank. In particular, approximation by low rank circulant matrices has important applications in signal and image processing. This paper discusses a procedure for using the fast Fourier transform to compute the nearest real circulant approximation C, with a specific rank, to a given real n-by-n matrix A. The low rank circulant real approximation problem is not as straightforward as the usual truncated singular value decomposition since a conjugate-even set of eigenvalues must be maintained to guarantee that C is real. A counter intuitive example is given.

Eigenvalue Approximations from Subspaces

Henk van der Vorst

Department of Mathematics
Utrecht University
The Netherlands
(vorst@math.ruu.nl)

Abstract

The computation of eigenvalues of high dimensional matrices, say of order 10,000 or more, is a very expensive task in terms of computer resources. An alternative technique is to restrict the matrix to convenient low-dimensional subspaces and to compute the eigenvalues of the restricted matrices. Krylov subspaces have become very popular for this purpose. The resulting eigenvalue approximations are known as the Ritz values.

We will give some background of this approach and we will discuss known and less well known aspects of the behavior of the eigenvalue approximations. There exists nice theory for the errors in the approximations for exact arithmetic. The behavior in finite precision arithmetic may lead to strange effects, as we will see.

In the 1990s much attention has been given to the so-called harmonic Ritz approximations. The harmonic Ritz pairs can be interpreted as to correspond with the inverse of A, restricted to A times the Krylov subspace. It turns out that one can generate a variety of such harmonic Ritz pairs, depending on a shift. The choice of the shift may be very critical for the detection of proper approximations. We will show some experimental results and surprising behavior.

The Definite Generalized Eigenvalue Problem:
A Reworked Perturbation Theory

Chi-Kwong Li and Roy Mathias

Department of Mathematics
College of William and Mary
Williamsburg, VA
(mathias@math.wm.edu)

Abstract

Let (A, B) be a definite pair of $n \times n$ Hermitian matrices. That is, $|x^*Ax| + |x^*Bx| \neq 0$ for all non-zero vectors $x \in C^n$. It is possible to find an $n \times n$ non-singular matrix X with unit columns such that

$$X^*(A + iB)X = \operatorname{diag}(\alpha_1 + i\beta_1, \ldots, \alpha_n + i\beta_n)$$

where α_j and β_j are real numbers. We call the pairs (α_j, β_j) *unit generalized eigenvalues* of the definite pair (A, B). These pairs have not been studied previously. We rework the perturbation theory for the eigenvalues and eigenvectors of the definite generalized eigenvalue problem $\beta Ax = \alpha Bx$ in terms of these unit generalized eigenvalues and show that they play a crucial rule in obtaining the best possible perturbation bounds. In particular, one can replace most instances of the Crawford number

$$c(A, B) = \min\{|x^*(A + iB)x| : x \in C^n, x^*x = 1\}$$

with the larger quantity

$$d_{\min} = \min\{|\alpha_j + i\beta_j| : j = 1, \ldots, n\}$$

in existing perturbation bounds. This results in bounds that can be stronger by an arbitrarily large factor. We also give a new measure of the separation of the jth eigenvalue from the kth:

$$|(\alpha_j + i\beta_j)\sin(\arg(\alpha_j + i\beta_j) - \arg(\alpha_k + i\beta_k)|.$$

This asymmetric measure is entirely new, and again results in bounds that can be arbitrarily stronger than the existing bounds. We show that all but one of our bounds are attainable.

In this talk I will present the results outlined above and highlight Pete's important insights and results in the area of the perturbation theory for the definite generalized eigenvalue problem.

Perturbation Analysis of the Generalized Singular Value Decomposition

Ji-guang Sun

Department of Computing Science
UmeåUniversity
Sweden
(jisun@cs.umu.se)

Abstract

The generalized singular value decomposition (GSVD) of two matrices A, B having the same number of columns is a very useful tool in many matrix computation problems. In this talk we present certain new results on perturbation analysis of the GSVD which include:

1. Computable formulas of condition numbers for a generalized singular value (GSV) of the matrix pair $\{A, B\}$;

2. A computable formula of the backward error of $\{A, B\}$ with respect to an approximate GSV and an associated approximate generalized singular vector group.

Certain connections between the new results for GSVD and the existing results in the special case of the ordinary singular value decomposition are made. Theoretical analysis and numerical examples show that the computable formulas of the condition numbers and backward error are useful tools for assessing the sensitivity of GSVs and the numerical quality of a computed GSVD.

Solution of Indefinite Systems and Inner/Outer Iterations

Gene H. Golub

Computer Science Department
Scientific Computing and Computational Mathematics Program
Stanford University
Stanford, CA 94305-9125
(golub@cssm.stanford.edu)

Abstract

We consider solving the linear system of equations $Ax = b$. In many situations, it is natural to rewrite the system as $Mx = Nx + b$. This leads to the iteration

$$Mx^{k+1} = Nx^k + b. \tag{1}$$

We can replace equation (1) by the equivalent system

$$Mz^k = r^k, \quad x^{k+1} = x^k + z^k, \tag{2}$$

where $r^k = b - Ax^k$. We have assumed that solving system (2) is easier than inverting A. Unfortunately, we may have to perform inner iterations to solve (2) so that we replace the algorithm described by (2) by

$$M\bar{z}^k = r^k + q^k, x^{k+1} = x^k + \bar{z}^k. \tag{3}$$

The vector q^k is the residual vector associated with the inner iteration; we continue to iterate until $\|q^k\| \leq \tau_k \|r^k\|$. The parameters $\{\tau_k\}$ play an important role in the convergence analysis, and we describe various strategies for choosing $\{\tau_k\}$. Our analysis will be independent of the method used for solving the system (3) and the inner iteration.

There are numerous problems that fit into the framework we have described. In particular, inner and outer iterations play an important role in domain decomposition, solution of non-symmetric systems where the coefficient matrix is real positive, and in the implementation of the method of Uzawa in solving the Stokes equation, whose discretization leads to an indefinite linear system of the KKT form. Indeed, we discuss at length a number of possible methods for solving KKT systems. Furthermore, we can apply these ideas to using Chebyshev polynomials for the acceleration of various iterative methods where fast direct methods are used for the inner iteration. Recently, we have analyzed with Qiang Ye the conjugate gradient method where the correction vector is not solved for exactly. We describe the results of various numerical experiments.

In addition, we shall discuss the solution of the generalized eigenvalue problem $Ax = \lambda Bx$. Here we shall show that it is not necessary to explicity invert or solve a system with the matrix B.

Bells and Whistles

Iain Duff

Rutherford Appleton Laboratory
Oxfordshire
United Kingdom
(I.Duff@rl.ac.uk)

Abstract

Iain Duff

We will start with a few reminiscences of Pete explaining our reason for the choice of title. The rest of the talk will be a technical one that will also respect the title. We will introduce a new code for the direct solution of sparse symmetric linear equations that solves indefinite systems with 2 x 2 pivoting for stability. This code, called MA57, is in HSL 2000 and supersedes the well used code MA27 in the Harwell Subroutine Library.

We will emphasize some of the novel features of this code and concentrate on the user interface so that we do not burden the listener with the gory details of a sparse implementation. Some of the features considered include start and restart facilities, matrix modification, partial solution for matrix factors, solution of multiple right-hand sides, and iterative refinement and error analysis. There are additional facilities within a Fortran 95 implementation that include the ability to identify and change pivots. Several of these facilities have been developed particularly to support optimization applications and the performance of the code on problems arising therefrom will be presented.

In Praise of the CS Decomposition

Nick Higham

Department of Mathematics
University of Manchester
United Kingdom
(higham@maths.man.ac.uk)

Abstract