Numerical Stability of Some Fast Algorithms for Structured Matrices

Dedicated to Gene Golub
on the occasion of his 65-th birthday

Richard P. Brent

Computer Sciences Laboratory
Australian National University
Canberra, ACT 0200, Australia

Richard.Brent@anu.edu.au

Abstract. We consider the numerical stability/instability of fast algorithms for solving systems of linear equations or linear least squares problems with a low displacement-rank structure. For example, the matrices involved may be Toeplitz or Hankel. In particular, we consider algorithms which incorporate pivoting without destroying the structure, such as the Gohberg-Kailath-Olshevsky (GKO) algorithm, and describe some recent results on the stability of these algorithms. We also compare these results with the corresponding stability results for algorithms based on the semi-normal equations and for the well known algorithms of Schur/Bareiss and Levinson.

1 Introduction

It is well known that systems of n linear equations with a low displacement rank (e.g. Toeplitz or Hankel matrices) can be solved in $O(n^2)$ arithmetic operations[1]. For positive definite Toeplitz matrices the first $O(n^2)$ algorithms were introduced by Kolmogorov [30], Wiener [43] and Levinson [31]. These algorithms are related to recursions of Szegö [40] for polynomials orthogonal on the unit circle. Another class of $O(n^2)$ algorithms, e.g. the Bareiss algorithm [2], are related to Schur's algorithm for finding the continued fraction representation of a holomorphic function in the unit disk [34]. This class can be generalized to cover unsymmetric matrices and more general "low displacement rank" matrices [28]. In this paper we consider the numerical stability of some of these algorithms. A more detailed survey is given in [29].

In the following, R denotes a structured matrix, T is a Toeplitz or Toeplitz-type matrix, P is a permutation matrix, L is lower triangular, U is upper triangular, and Q is orthogonal. In error bounds, $O_n(\varepsilon)$ means $O(\varepsilon f(n))$, where $f(n)$ is a polynomial in n.

[1] Asymptotically faster algorithms exist [1, 8], but are not considered here.

rpb173 typeset using LaTeX

2 Classes of Structured Matrices

Structured matrices R satisfy a *Sylvester equation* which has the form

$$\nabla_{\{A_f,A_b\}}(R) = A_f R - R A_b = \Phi\Psi \ , \tag{1}$$

where A_f and A_b have some simple structure (usually banded, with 3 or fewer full diagonals), Φ and Ψ are $n \times \alpha$ and $\alpha \times n$ respectively, and α is some fixed integer. The pair of matrices (Φ, Ψ) is called the $\{A_f, A_b\}$*-generator* of R.

α is called the $\{A_f, A_b\}$*-displacement rank* of R. We are interested in cases where α is small (say at most 4).

Cauchy matrices

Particular choices of A_f and A_b lead to definitions of basic classes of matrices. Thus, for a Cauchy matrix

$$C(\mathbf{t}, \mathbf{s}) = \left[\frac{1}{t_i - s_j} \right]_{ij} \ ,$$

we have

$$A_f = D_t = \mathrm{diag}(t_1, t_2, \ldots, t_n) \ ,$$
$$A_b = D_s = \mathrm{diag}(s_1, s_2, \ldots, s_n)$$

and

$$\Phi^T = \Psi = [1, 1, \ldots, 1] \ .$$

More general matrices, where Φ and Ψ are any rank-α matrices, are called *Cauchy-type.*

Toeplitz matrices

For a Toeplitz matrix $T = [t_{ij}] = [a_{i-j}]$, we take

$$A_f = Z_1 = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & & & 0 \\ 0 & 1 & & & \vdots \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}, \quad A_b = Z_{-1} = \begin{bmatrix} 0 & 0 & \cdots & 0 & -1 \\ 1 & 0 & & & 0 \\ 0 & 1 & & & \vdots \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix},$$

$$\Phi = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ a_0 & a_{1-n} + a_1 & \cdots & a_{-1} + a_{n-1} \end{bmatrix}^T,$$

and

$$\Psi = \begin{bmatrix} a_{n-1} - a_{-1} & \cdots & a_1 - a_{1-n} & a_0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} \ .$$

We can generalize to *Toeplitz-type* matrices by taking Φ and Ψ to be general rank-α matrices.

3 Structured Gaussian Elimination

Let an input matrix, R_1, have the partitioning

$$R_1 = \begin{bmatrix} d_1 & \mathbf{w}_1^T \\ \mathbf{y}_1 & \tilde{R}_1 \end{bmatrix} .$$

The first step of normal Gaussian elimination is to premultiply R_1 by

$$\begin{bmatrix} 1 & \mathbf{0}^T \\ -\mathbf{y}_1/d_1 & I \end{bmatrix} ,$$

which reduces it to

$$\begin{bmatrix} d_1 & \mathbf{w}_1^T \\ \mathbf{0} & R_2 \end{bmatrix} ,$$

where

$$R_2 = \tilde{R}_1 - \mathbf{y}_1\mathbf{w}_1^T/d_1$$

is the *Schur complement* of d_1 in R_1. At this stage, R_1 has the factorization

$$R_1 = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{y}_1/d_1 & I \end{bmatrix} \begin{bmatrix} d_1 & \mathbf{w}_1^T \\ \mathbf{0} & R_2 \end{bmatrix} .$$

One can proceed recursively with the Schur complement R_2, eventually obtaining a factorization $R_1 = LU$.

The key to *structured* Gaussian elimination is the fact that the displacement structure is preserved under Schur complementation, and that the generators for the Schur complement R_{k+1} can be computed from the generators of R_k in $O(n)$ operations.

Row and/or column interchanges destroy the structure of matrices such as Toeplitz matrices. However, if A_f is diagonal (which is the case for Cauchy and Vandermonde type matrices), then *the structure is preserved under row permutations.*

This observation leads to the *GKO-Cauchy* algorithm [21] for fast factorization of Cauchy-type matrices with partial pivoting, and many recent variations on the theme by Boros, Gohberg, Ming Gu, Heinig, Kailath, Olshevsky, M. Stewart, *et al*: see [7, 21, 23, 26, 35].

The GKO-Toeplitz algorithm

Heinig [26] showed that, if T is a Toeplitz-type matrix, then

$$R = FTD^{-1}F^*$$

is a Cauchy-type matrix, where

$$F = \frac{1}{\sqrt{n}}[e^{2\pi i(k-1)(j-1)/n}]_{1\le k,j\le n}$$

is the Discrete Fourier Transform matrix,

$$D = \mathrm{diag}(1, e^{\pi i/n}, \ldots, e^{\pi i(n-1)/n}),$$

and the generators of T and R are simply related.

The transformation $T \leftrightarrow R$ is perfectly stable because F and D are unitary. Note that R is (in general) complex even if T is real.

Heinig's observation was exploited by Gohberg, Kailath and Olshevsky [21]: R can be factorized as $R = P^T LU$ using GKO-Cauchy. Thus, from the factorization

$$T = F^* P^T LUFD \ ,$$

a linear system involving T can be solved in $O(n^2)$ operations. The full procedure of conversion to Cauchy form, factorization, and solution requires $O(n^2)$ (complex) operations.

Other structured matrices, such as Hankel, Toeplitz-plus-Hankel, Vandermonde, Chebyshev-Vandermonde, etc, can be converted to Cauchy-type matrices in a similar way.

Error Analysis

Because GKO-Cauchy and GKO-Toeplitz involve partial pivoting, we might guess that their stability would be similar to that of Gaussian elimination with partial pivoting. Unfortunately, there is a flaw in this reasoning. During GKO-Cauchy the *generators* have to be transformed, and the partial pivoting does not ensure that the transformed generators are small.

Sweet and Brent [39] show that significant generator growth can occur if all the elements of $\Phi\Psi$ are small compared to those of $|\Phi||\Psi|$. This can not happen for ordinary Cauchy matrices because $\Phi^{(k)}$ and $\Psi^{(k)}$ have only one column and one row respectively. However, it can happen for higher displacement-rank Cauchy-type matrices, even if the original matrix is well-conditioned.

The Toeplitz Case

In the Toeplitz case there is an extra constraint on the selection of Φ and Ψ, but it is still possible to give examples where the normalized solution error grows like κ^2 and the normalized residual grows like κ, where κ is the condition number of the Toeplitz matrix. Thus, the GKO-Toeplitz algorithm is (at best) weakly stable[2].

It is easy to think of modified algorithms which avoid the examples given by Sweet and Brent, but it is difficult to prove that they are stable in all cases. Stability depends on the worst case, which may be rare and hard to find by random sampling.

The problem with the original GKO algorithm is growth in the generators. Ming Gu suggested exploiting the fact that the generators are not unique. Recall

[2] For definitions of stability and weak stability, see [5, 9, 10].

the Sylvester equation (1). Clearly we can replace Φ by ΦM and Ψ by $M^{-1}\Psi$, where M is any invertible $\alpha \times \alpha$ matrix, because this does not change the product $\Phi\Psi$. Similarly at later stages of the GKO algorithm.

Ming Gu [23] proposes taking M to orthogonalize the columns of Φ (that is, at each stage we do an orthogonal factorization of the generators). Michael Stewart [35] proposes a (cheaper) LU factorization of the generators. In both cases, clever pivoting schemes give error bounds analogous to those for Gaussian elimination with partial pivoting.

Gu and Stewart's error bounds

The error bounds obtained by Ming Gu and Michael Stewart involve a factor K^n, where K depends on the ratio of the largest to smallest modulus elements in the Cauchy matrix

$$\left[\frac{1}{t_i - s_j}\right]_{ij} .$$

Although this is unsatisfactory, it is similar to the factor 2^{n-1} in the error bound for Gaussian elimination with partial pivoting.

Michael Stewart [35] gives some interesting numerical results which indicate that his scheme works well, but more numerical experience is necessary before a definite conclusion can be reached.

In practice, we can use an $O(n^2)$ algorithm such as Michael Stewart's, check the residual, and resort to iterative refinement or a stable $O(n^3)$ algorithm in the (rare) cases that it is necessary.

4 Positive Definite Structured Matrices

An important class of algorithms, typified by the algorithm of Bareiss [2], find an LU factorization of a Toeplitz matrix T, and (in the symmetric case) are related to the classical algorithm of Schur [20, 34].

It is interesting to consider the numerical properties of these algorithms and compare with the numerical properties of the Levinson algorithm (which essentially finds an LU factorization of T^{-1}).

The Bareiss algorithm for positive definite matrices

Bojanczyk, Brent, de Hoog and Sweet[3] [6, 37] have shown that the numerical properties of the Bareiss algorithm are similar to those of Gaussian elimination (*without* pivoting). Thus, the algorithm is stable for positive definite symmetric Toeplitz matrices.

The Levinson algorithm can be shown to be weakly stable for bounded n, and numerical results by Varah [42], BBHS and others suggest that this is all

[3] Abbreviated BBHS.

that we can expect. Thus, the Bareiss algorithm is (generally) better numerically than the Levinson algorithm.

Cybenko [15] showed that if certain quantities called "reflection coefficients" are positive then the Levinson-Durbin algorithm for solving the Yule-Walker equations (a positive-definite system with special right-hand side) is stable. However, "random" positive-definite Toeplitz matrices do not usually satisfy Cybenko's condition.

The generalized Schur algorithm

The Schur algorithm can be generalized to factor a large variety of structured matrices – see Kailath *et al* [27, 28]. For example, the generalized Schur algorithm applies to block Toeplitz matrices, Toeplitz block matrices, and to matrices of the form $T^T T$, where T is rectangular Toeplitz.

It is natural to ask if the stability results of BBHS (which are for the classical Schur/Bareiss algorithm) extend to the generalized Schur algorithm. This was considered by M. Stewart and Van Dooren [36] and by Chandrasekharan and Sayed [12]. (The results were obtained independently by the two pairs of authors, and the "generalized Schur algorithm" considered in each case is slightly different – for details see [29].)

The conclusion is that the generalized Schur algorithm is stable for positive definite symmetric (or Hermitian) matrices, provided that the hyperbolic transformations in the algorithm are implemented correctly. In contrast, BBHS showed that stability of the classical Schur/Bareiss algorithm is not so dependent on details of the implementation.

5 Fast Orthogonal Factorization

In an attempt to achieve stability without pivoting, and to solve $m \times n$ least squares problems ($m \geq n$), it is natural to consider algorithms for computing an orthogonal factorization

$$T = QU$$

of T. The first such $O(n^2)$ algorithm[4] was introduced by Sweet [37]. Unfortunately, Sweet's algorithm is unstable: it depends on the condition of a submatrix of T – see Luk and Qiao [32].

Other $O(n^2)$ algorithms for computing the matrices Q and U or U^{-1} were given by Bojanczyk, Brent and de Hoog[5] [4], Chun *et al* [14], Cybenko [16], and Qiao [33], but none of them has been shown to be stable, and in several cases examples show that they are unstable.

Unlike the classical $O(n^3)$ Givens or Householder algorithms, the $O(n^2)$ algorithms do not form Q in a numerically stable manner as a product of matrices which are (close to) orthogonal.

[4] More precisely, $O(mn)$. For simplicity, in the time bounds we assume $m = O(n)$.
[5] Abbreviated BBH.

For example, the algorithms of Bojanczyk, Brent and de Hoog [4] and Chun *et al* [14] depend on Cholesky downdating, and numerical experiments show that they do not give a Q which is close to orthogonal.

The generalized Schur algorithm, applied to T^TT, computes the upper triangular matrix U but not the orthogonal matrix Q.

Use of the semi-normal equations

It can be shown that, provided the Cholesky downdates are implemented in a certain way (analogous to the condition for the stability of the generalized Schur algorithm), the BBH algorithm computes U in a weakly stable manner [5]. In fact, the computed upper triangular matrix $\tilde{U}$ is about as good as can be obtained by performing a Cholesky factorization of T^TT, so

$$\|T^TT - \tilde{U}^T\tilde{U}\| / \|T^TT\| = O_m(\varepsilon) \ .$$

Thus, by solving

$$\tilde{U}^T\tilde{U}x = T^Tb$$

(the so-called *semi-normal* equations) we have a *weakly stable* algorithm for the solution of general Toeplitz systems $Tx = b$ in $O(n^2)$ operations. The solution can be improved by iterative refinement if desired. Note that the computation of Q is avoided, and the algorithm is applicable to full-rank Toeplitz least squares problems.

Computing Q stably

It is difficult to give a satisfactory $O(n^2)$ algorithm for the computation of Q in the factorization

$$T = QU \tag{2}$$

Chandrasekharan and Sayed give a stable algorithm to compute the factorization

$$T = LQU \tag{3}$$

where L is lower triangular. Their algorithm can be used to solve linear equations, but not least squares problems, because T has to be square, and in any case the matrix Q in (3) is different from the matrix Q in (2). Because their algorithm involves embedding the $n \times n$ matrix T in a $2n \times 2n$ matrix

$$\begin{bmatrix} T^TT & T^T \\ T & 0 \end{bmatrix} ,$$

the constant factors in the operation count are large: $59n^2 + O(n \log n)$, which should be compared to $8n^2 + O(n \log n)$ for BBH and the semi-normal equations.

References

1. G. S. Ammar and W. B. Gragg, "Superfast solution of real positive definite Toeplitz systems", *SIMAX* 9 (1988), 61–76.
2. E. H. Bareiss, "Numerical solution of linear equations with Toeplitz and vector Toeplitz matrices", *Numer. Math.* 13 (1969), 404–424.
3. A. W. Bojanczyk, R. P. Brent, P. Van Dooren and F. R. de Hoog, "A note on downdating the Cholesky factorization", *SISSC* 8 (1987), 210–220.
4. A. W. Bojanczyk, R. P. Brent and F. R. de Hoog, "QR factorization of Toeplitz matrices", *Numer. Math.* 49 (1986), 81–94.
5. A. W. Bojanczyk, R. P. Brent and F. R. de Hoog, "Stability analysis of a general Toeplitz systems solver", *Numerical Algorithms* 10 (1995), 225–244.
6. A. W. Bojanczyk, R. P. Brent, F. R. de Hoog and D. R. Sweet, "On the stability of the Bareiss and related Toeplitz factorization algorithms", *SIAM J. Matrix Anal. Appl.* 16 (1995), 40–57.
7. T. Boros, T. Kailath and V. Olshevsky, "Fast algorithms for solving Cauchy linear systems", preprint, 1995.
8. R. P. Brent, F. G. Gustavson and D. Y. Y. Yun, "Fast solution of Toeplitz systems of equations and computation of Padé approximants", *J. Algorithms* 1 (1980), 259–295.
9. J. R. Bunch, "Stability of methods for solving Toeplitz systems of equations", *SISSC* 6 (1985), 349–364.
10. J. R. Bunch, "The weak and strong stability of algorithms in numerical linear algebra", *Linear Alg. Appl.* 88/89 (1987), 49–66.
11. J. R. Bunch, "Matrix properties of the Levinson and Schur algorithms", *J. Numerical Linear Algebra with Applications* 1 (1992), 183–198.
12. S. Chandrasekaran and A. H. Sayed, *Stabilizing the generalized Schur algorithm*, *SIMAX* 17 (1996), 950–983.
13. S. Chandrasekaran and A. H. Sayed, *A fast stable solver for nonsymmetric Toeplitz and quasi-Toeplitz systems of linear equations*, preprint, Jan. 1996.
14. J. Chun, T. Kailath and H. Lev-Ari, "Fast parallel algorithms for QR and triangular factorization", *SIAM J. Sci. Stat. Computing* 8 (1987), 899–913.
15. G. Cybenko, "The numerical stability of the Levinson-Durbin algorithm for Toeplitz systems of equations", *SIAM J. Sci. Stat. Computing* 1 (1980), 303–319.
16. G. Cybenko, "Fast Toeplitz orthogonalization using inner products", *SIAM J. Sci. Stat. Computing* 8 (1987), 734–740.
17. J.-M. Delosme and I. C. F. Ipsen, "Parallel solution of symmetric positive definite systems with hyperbolic rotations", *Linear Algebra Appl.* 77 (1986), 75–111.
18. P. Delsarte, Y. V. Genin and Y. G. Kamp, "A generalisation of the Levinson algorithm for Hermitian Toeplitz matrices with any rank profile", *IEEE Trans. Acoustics, Speech and Signal Processing* ASSP–33 (1985), 964–971.
19. J. Durbin, "The fitting of time-series models", *Rev. Int. Stat. Inst.* 28 (1959), 229–249.
20. I. Gohberg (editor), *I. Schur Methods in Operator Theory and Signal Processing* (Operator Theory: Advances and Applications, Volume 18), Birkhäuser Verlag, Basel, 1986.
21. I. Gohberg, T. Kailath and V. Olshevsky, "Gaussian elimination with partial pivoting for matrices with displacement structure", *Math. Comp.* 64 (1995), 1557–1576.
22. G. H. Golub and C. Van Loan, *Matrix Computations*, second edition, Johns Hopkins Press, Baltimore, Maryland, 1989.

23. Ming Gu, *Stable and efficient algorithms for structured systems of linear equations*, Tech. Report LBL-37690, Lawrence Berkeley Laboratory, Aug. 1995.
24. Ming Gu, *New fast algorithms for structured least squares problems*, Tech. Report LBL-37878, Lawrence Berkeley Laboratory, Nov. 1995.
25. P. C. Hansen and H. Gesmar, "Fast orthogonal decomposition of rank deficient Toeplitz matrices", *Numerical Algorithms* 4 (1993), 151–166.
26. G. Heinig, "Inversion of generalized Cauchy matrices and other classes of structured matrices", *Linear Algebra for Signal Processing, IMA Volumes in Mathematics and its Applications, Vol. 69,* Springer, 1994, 95–114.
27. T. Kailath and J. Chun, "Generalized displacement structure for block-Toeplitz, Toeplitz-block, and Toeplitz-derived matrices", *SIAM J. Matrix Anal. Appl.* 15 (1994), 114–128.
28. T. Kailath and A. H. Sayed, "Displacement structure: theory and applications", *SIAM Review* 37 (1995), 297–386.
29. T. Kailath and A. H. Sayed (editors), *Fast Reliable Algorithms for Matrices with Structure*, to appear.
30. A. N. Kolmogorov, "Interpolation and extrapolation of stationary random sequences", *Izvestia Akad. Nauk SSSR* 5 (1941), 3–11 (in Russian). German summary, *ibid* 11–14.
31. N. Levinson, "The Wiener RMS (Root-Mean-Square) error criterion in filter design and prediction", *J. Math. Phys.* 25 (1947), 261–278.
32. F. T. Luk and S. Qiao, "A fast but unstable orthogonal triangularization technique for Toeplitz matrices", *Linear Algebra Appl.* 88/89 (1987), 495–506.
33. S. Qiao, "Hybrid algorithm for fast Toeplitz orthogonalization", *Numer. Math.* 53 (1988), 351–366.
34. I. Schur, "Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind", *J. fur die Reine und Angewandte Mathematik* 147 (1917), 205–232. English translation in [20], 31–59.
35. M. Stewart, "Stable pivoting for the fast factorization of Cauchy-like matrices", preprint, Jan. 13, 1997.
36. M. Stewart and P. Van Dooren, "Stability issues in the factorization of structured matrices", preprint, Aug. 1996. To appear in *SIMAX*.
37. D. R. Sweet, *Numerical Methods for Toeplitz Matrices*, PhD thesis, University of Adelaide, 1982.
38. D. R. Sweet, "Fast Toeplitz orthogonalization", *Numer. Math.* 43 (1984), 1–21.
39. D. R. Sweet and R. P. Brent, "Error analysis of a fast partial pivoting method for structured matrices", *Proceedings SPIE, Volume 2563, Advanced Signal Processing Algorithms* SPIE, Bellingham, Washington, 1995, 266–280.
40. G. Szegö, *Orthogonal Polynomials*, AMS Colloquium publ. XXIII, AMS, Providence, Rhode Island, 1939.
41. W. F. Trench, "An algorithm for the inversion of finite Toeplitz matrices", *J. SIAM (SIAM J. Appl. Math.)* 12 (1964), 515–522.
42. J. M. Varah, "The prolate matrix", *Linear Algebra Appl.* 187 (1993), 269–278.
43. N. Wiener, *Extrapolation, Interpolation and Smoothing of Stationary Time Series, with Engineering Applications*, Technology Press and Wiley, New York, 1949 (originally published in 1941 as a Technical Report).