Fast transforms: Banded matrices with banded inverses

Fast transforms: Banded matrices with banded inverses

Gilbert Strang[1]

Massachusetts Institute of Technology, Cambridge, MA 02139

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected in 2009.

Contributed by Gilbert Strang, April 22, 2010 (sent for review March 9, 2010)

It is unusual for both A and A^{-1} to be banded—but this can be a valuable property in applications. Block-diagonal matrices F are the simplest examples; wavelet transforms are more subtle. We show that every example can be factored into $A = F_1 \cdots F_N$ where N is controlled by the bandwidths of A and A^{-1} (but not by their size, so this extends to infinite matrices and leads to new matrix groups).

Bruhat | CMV matrix | factorization | wavelet | permutation

1. Introduction

An invertible transform $y = Ax$ expresses the vector x in a new basis. The inverse transform $x = A^{-1}y$ reconstructs x as a combination of the basis vectors with coefficients from y. This matrix-vector multiplication $A^{-1}y$ is best seen as a combination of the columns of A^{-1} (which are the basis vectors).

In applications, the transform is often intended to separate signal from noise, important information from unimportant. This separation is followed by a divorce, when the full vector y is compressed to $\hat{y}$.

This compressed form $\hat{y}$ is all we keep in the reconstruction step:

$$\text{Input signal } x \to y = Ax \to \hat{y} \to \hat{x} = A^{-1}\hat{y} \text{ Compressed signal.}$$

The choice of transform (the choice of a good basis) is crucial. It must serve its purpose, and it must be fast. This paper is concerned most of all with speed.

The Fourier transform is nearly perfect, when noise is identified with high frequencies. Quick execution of A and A^{-1} comes from the Fast Fourier Transform. And yet, we want to go further. Clarity in the frequency domain means obscurity in the time domain. For signals that change quickly, a "short time" transform is needed.

This leads to banded matrices. The entries are $a_{ij} = 0$ when $|i - j| > w$. Nonzero entries are in a band along the main diagonal. Then A acts locally and each y_k comes quickly from $x_{k-w}, \ldots x_{k+w}$. The challenge is to achieve a banded A^{-1}, so the inverse transform is also fast. Typical band matrices have full inverses, and the exceptions to this rule are the subject of this paper.

Briefly, we want to factor A in a way that makes the property of a banded inverse evident. The factors F will be block diagonal (with invertible blocks, 2 by 2 or 1 by 1). Then $F_1 \cdots F_N$ is clearly banded with banded inverse.

We recall two constructions of A that are successful and useful:

1. Block Toeplitz matrices in which the matrix polynomial like $M(z) = R + Sz + Tz^2$ has a monomial determinant $= cz^k$:

$$A = \begin{bmatrix} \ddots & \ddots & & & \\ R & S & T & & \\ & R & S & T & \\ & & \ddots & \ddots & \ddots \end{bmatrix} \quad \text{(often doubly infinite).}$$

2. "CMV matrices" with 2 by 2 singular blocks P_i, Q_i (r is 1 by 2):

$$A = \begin{bmatrix} r & & & \\ P_1 & Q_1 & & \\ & P_2 & Q_2 & \\ & & \ddots & \ddots \end{bmatrix} \quad \text{with rank } (P_i) = \text{rank } (Q_i) = 1.$$

Construction 1 is a "filter bank," time invariant because A is block Toeplitz. Suitable choices of the filters $R,S,T,\ldots$ lead to wavelets (1, 2).

Construction 2 produced new families of orthogonal polynomials on the circle $|z| = 1$ (3). Here we drop the requirement that A is orthogonal; the bandedness of A^{-1} is the important point.

For CMV matrices, the factors in $A = F_1F_2$ are known. For block Toeplitz matrices, the factorization of matrix functions $M(z)$ has an overwhelming history (including Wiener–Hopf). By combining the two, the CMV construction extends to more blocks per row. These "time-varying filter banks" may find application in signal processing, using the factorization.

Beyond 1 and 2, our true goal is to factor all banded matrices with banded inverses. Permutation matrices are a third example, when no entry is more than w positions out of place. (Then each 2 by 2 block in each factor F executes a transposition of neighbors.) Our main result is that a factorization of this kind is always possible, in which the number of factors in $F_1 \cdots F_N$ is controlled by the bandwidths of A and A^{-1}.

Theorem. *Suppose A has bandwidth w and A^{-1} has bandwidth W. Then A is a product $F_1 \cdots F_N$ of block-diagonal factors. Each F is composed of 2 by 2 and 1 by 1 blocks, with $N \le C(w^3 + W^3)$.*

We have not minimized N (but the low numbers $N = 2$ for CMV and $N = L$ for L blocks per row are potentially important). The essential point is that N is independent of the matrix dimension n. In the proof, the bandedness of A^{-1} implies rank conditions on certain submatrices of A. Together with the bandedness of A itself, these conditions yield the factorization into Fs.

The independence from n suggests an entirely algebraic statement for $n = \infty$: The invertible block-diagonal matrices F generate the group of singly infinite banded matrices with banded inverses.

2. Factorizations in Constructions 1 and 2.

Before the general case of banded A and A^{-1}, allow us to develop the block-matrix factorizations. If A starts with L blocks per row, the first step is to reach factors with two blocks per row.

For our block Toeplitz A with $L = 3$, we remove from $M(z) = R + Sz + Tz^2$ a linear factor $P + Qz$ of a special form. The 2 by 2 matrices P and Q are complementary projections on column vectors r and t:

Author contributions: G.S. performed research and wrote the paper.

The authors declare no conflict of interest.

[1]E-mail: gs@math.mit.edu.

$$Pr = r, \quad Pt = 0, \quad Qr = 0, \quad Qt = t,$$
$$PQ = QP = 0, \quad P + Q = I.$$

Knowing those eigenvalues 1 and 0, it follows directly that

$$(P + Qz)r = r, \quad (P + Qz)t = zt, \quad \text{and}$$
$$(P + Qz)^{-1} = P + Qz^{-1}.$$

We want to factor $M(z) = R + Sz + Tz^2$ when its determinant is a monomial. It is this property that makes A^{-1} banded. For doubly infinite Toeplitz matrices, A^{-1} is associated with the matrix function $(M(z))^{-1}$. That inverse requires division by the determinant. So $(M(z))^{-1}$ is a polynomial (in z and z^{-1}) only when $\det(M(z))$ is a monomial cz^k. Polynomials correspond to banded matrices.

This could not be achieved in the scalar case, where the inverse of $1 - \frac{1}{2}z$ (for example) is an infinite series. The inverse of the Toeplitz matrix with diagonals 1 and $-\frac{1}{2}$ will have diagonals $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$ from $1/(1 - \frac{1}{2}z)$. This inverse matrix is not banded —the reciprocal of a scalar polynomial is not a polynomial. It is the extra flexibility in the block-matrix case that allows $\det M$ to be a monomial and A^{-1} to be banded. This leads to wavelets.

To factor $R + Sz + Tz^2$ when R and T have rank 1, choose r and t in the ranges of R and T. (More explicitly, $R = ru^{\mathrm{T}}$ and $T = tv^{\mathrm{T}}$ and S turns out to be a combination of rv^{T} and tu^{T}. We assume r is not parallel to t; degenerate cases are not included here.) Then $PT = 0$ and $QR = 0$ give the required separation of the quadratic $M(z)$ into linear factors:

$$(P + Qz)^{-1}(R + Sz + Tz^2) = (P + Qz^{-1})(R + Sz + Tz^2)$$
$$= (PR + QS) + (PS + QT)z.$$

The infinite Toeplitz matrix with blocks P, Q on two diagonals times the new one with blocks $PR + QS$ and $PS + QT$ equals A (with three blocks R, S, T). Because A and the P, Q matrix have banded inverse, so does the new one. The two new blocks will then have rank 1—also as a consequence of the z and z^3 zero terms in $\det(M(z))$.

Of course far more general factorizations of matrix functions come from the classical theory. Their form is $M = M_+ D M_-$ where D is a diagonal block of monomials (those powers are the "partial indices" of $M(z)$). After G. D. Birkhoff's proof, he gave precedence to Hilbert and to Plemelj. For singly infinite systems $Ax = b$, the Wiener–Hopf factorization $A = UL$ (and not LU) preserves the Toeplitz structure. Recent progress was led by Gohberg, and the exposition in ref. 4 is particularly clear.

In the time-varying non-Toeplitz case, blocks R_i, S_i, and T_i in row i will interact with neighboring rows in $A^{-1}A = I$. Inverting $M(z)$ or $M_i(z)$ no longer yields A^{-1}. But a matrix multiplication parallel to the symbolic one displayed above shows how to construct the complementary projections P_i and Q_i:

$$\begin{bmatrix} \cdot & & & \\ Q_i & P_i & & \\ & Q_{i+1} & P_{i+1} & \\ & & \cdot & \cdot \end{bmatrix} \begin{bmatrix} R_{i-1} & S_{i-1} & T_{i-1} & \\ & R_i & S_i & T_i \\ & & \cdot & \cdot \\ & & & \cdot \end{bmatrix}$$

is to be bidiagonal. The outer blocks along row i are $Q_i R_{i-1}$ and $P_i T_i$. Those are zero if P_i and Q_i are complementary projections on the ranges of R_{i-1} and T_i, which must be rank-one blocks if the complete R, S, T matrix has banded inverse. (Again we treat only the nondegenerate case.) This rank-one requirement comes from the rank conditions for a banded inverse, described below in the proof of our general theorem.

The conclusion is that if the matrix A with L blocks per row (2 by 2 matrices R_i,S_i,…) has a banded inverse, then A can be factored recursively into a product of $L - 1$ block bidiagonal matrices with rank-one blocks.

Already in the Toeplitz case, it is not required that each pair of rows has an equal number of nonzeros. When the eight nonzeros from the two blocks are the celebrated Daubechies coefficients, F_1 and F_2 were computed in ref. 5. But there are 3–5 and 2–6 wavelet transforms as well as that 4–4 choice (and 9–7 is one of the best). In all these cases time-varying versions become feasible by using the new factorization.

Now we recall the final step to the Fs. Start with 2 by 2 blocks P_i and Q_i. Then F_1 and F_2 have 2 by 2 blocks along their diagonals, and the key is that the blocks in F_2 are offset from the blocks in F_1.

$$F_1 = \begin{bmatrix} 1 & & & \\ & \begin{bmatrix} c_2 & c_3 \end{bmatrix} & & \\ & & \begin{bmatrix} c_4 & c_5 \end{bmatrix} & \end{bmatrix} \qquad F_2 = \begin{bmatrix} \begin{bmatrix} d_1^T \\ d_2^T \end{bmatrix} & & \\ & \begin{bmatrix} d_3^T \\ d_4^T \end{bmatrix} & \\ & & \cdot \end{bmatrix}$$

In F_1F_2, we multiply columns of F_1 times corresponding rows of F_2 (and add those rank-one products). Because of the offset, c_2 multiplies the second row d_2^{T} of a block in F_2, and c_3 multiplies the first row d_3^{T} of the following block, to produce the P and Q blocks in the CMV matrix $A = F_1F_2$:

$$c_2 d_2^{\mathrm{T}} = P_1 \quad \text{and} \quad c_3 d_3^{\mathrm{T}} = Q_1.$$

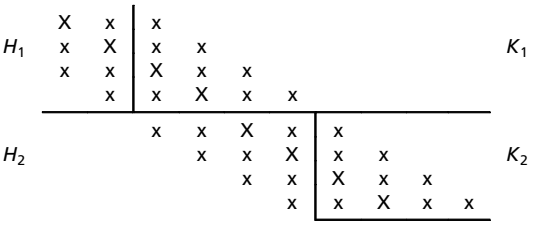
This simple but revolutionary idea appeared in refs. 3 and 6. Without an offset (a time delay), F_1F_2 would be block diagonal again.

Multiplying two CMV matrices F_1F_2 and F_3F_4, we arrange for F_3 to share the form of F_2. The overall product $F_1(F_2F_3)F_4$ has 3 factors from $L = 3$ blocks per row. Boundary rows are included in the general theorem that we now prove.

3. Banded Matrices with Banded Inverses

To prove the theorem, A is made diagonal by a sequence of elimination steps. We describe those steps when A and A^{-1} have bandwidth $w = W = 2$. The consequence of bandwidth W for A^{-1} is crucial: All submatrices H and K of A, below the Wth upper diagonal and above the Wth lower diagonal, have rank $\leq W$. This rank condition was discovered by Asplund (7), and the expository paper (8) describes two proofs.

Proof of Theorem. Elimination begins on the first $2W = 4$ rows of A. Separate those rows into submatrices H_1 and K_1 with ranks $\leq W$, as shown. They must each have rank equal to $W = 2$, because those four rows of A are independent (A is invertible). Possible nonzeros are marked by x and X, with $w = 2$.



The usual row operations will produce nonzeros in the first two diagonal positions X and zeros elsewhere in H_1. Because rows 3 and 4 of A remain independent, the new rows 3 and 4 of K_1 must be independent. Row operations will remove rows 1 and 2 of K_1, without changing H_1. Then operations on columns 3, 4, 5, and 6 of A will produce nonzeros in the two positions X and zeros elsewhere in K_1.

Now move to rows 5–8. Again H_2 and K_2 must have rank exactly $W = 2$. The last two columns of (the new) H_2 must be independent, because those columns are now zero in K_1. Use these columns to produce zeros in the first two columns of H_2, without changing K_1. Now H_2 and K_2 are in exactly the same situation as the original H_1 and K_1.

Operate as before on rows 5–8 of the current A and on columns 7–10, to produce nonzeros only in the diagonal positions X. All row operations are left multiplications by elimination matrices (which we factor below into products of our admissible Fs.) The key point is that an operation on rows 1–4 and an operation on rows 5–8 can be carried out together (in parallel by the same Fs). Similarly, column operations on A are right multiplications, and columns 3–6 can be changed in parallel with columns 7–10.

Conclusion. A block-diagonal matrix B_r acts on $2W = 4$ rows at a time. A block-diagonal matrix B_c acts on the columns, and B_rAB_c is diagonal. Note that B_c is offset so it starts on columns 3–6.

The reader might consider a permutation matrix with two nonzeros in each submatrix H_1, K_1, H_2, K_2 .... Exchanges of neighboring rows and of neighboring columns will produce I. The row exchanges within $H_1, H_2, \ldots$ and the column exchanges within $K_1, K_2, \ldots$ can be carried out in parallel.

B_r and B_c can be executed by block-diagonal Fs. The usual elimination steps subtract a multiple of one row or column from another (within the Hs and the Ks). Each operation can be achieved by a product of $2d - 1$ factors F, when the "x" to be removed is d rows or columns away from the "1." (For $d = 1$, F is a Gauss matrix G with a single nonzero next to the diagonal of 1s. For $d = 2$, the product $F_1F_2F_3 = PGP$ moves that nonzero by one position when the Ps exchange rows and exchange columns.) Certainly W^3 factors will suffice to remove all the off-diagonal nonzeros in each H and K.

This completes the proof when $w \le W$. In case $w > W$, we operate instead on the matrix A^{-1}. Its blocks H and K will have $2w$ rows. Again, it reduces to an invertible diagonal matrix X by B_r and B_c, and those factor into Fs. The number of Fs is independent of n, and their construction still succeeds for $n = \infty$.

4. Permutation Matrices

An n by n permutation matrix P has a single 1 in each row and each column. Let the column number for that nonzero entry in row i be $p(i)$. Then $p = (p(1), \ldots, p(n))$ is the permutation of $1, \ldots, n$ associated with P. The bandwidth w of both P and $P^{-1} = P^{\mathrm{T}}$ is the largest distance $|i - p(i)|$.

The block-diagonal factors in $P = F_1 \cdots F_N$ can also be permutation matrices. Because the diagonal blocks in F are 1 by 1 or 2 by 2, the associated permutation can only exchange disjoint pairs of neighbors. A sequence of $N = 5$ F's acts on the example $p = (4,5,6,1,2,3)$ to produce the identity permutation:

$$456123 \to 451623 \to 415263 \to 142536 \to 124356 \to 123456.$$

The original P from 456123 had bandwidth $w = 3$. Every entry must move 3 positions. The 6 by 6 matrix is $P = [0\,I\,;\,I\,0]$ with 3 by 3 blocks. The number of "disjoint exchanges in parallel" was $N = 5 = 2w - 1$. We believe that this example is extreme, and we conjecture that $2w - 1$ factors F are always sufficient for any P.

The algorithm itself, forced by the limitation of 2 by 2 diagonal blocks in each F, is a "parallel bubblesort." This problem of sorting has had enormous attention in computer science. It is fascinating that ordinary bubblesort is completely out of favor. But the parallel version seems appropriate here, and our $2w - 1$ conjecture was waiting for the right idea. Two proofs have now been found, by Panova and by Albert, Li, and Yu.

5. Infinite Matrices

For a banded infinite matrix, the elimination steps still succeed. When the inverse is also banded, the central ideas of linear algebra are undisturbed. There are no issues of convergence because all vectors have finitely many nonzeros.

But the crucial matrix theorem needed for this paper was hidden in Section 3 above. For a matrix with bandwidth W, all submatrices H below diagonal W of the inverse matrix have rank $\le W$. In our application the banded matrix was A^{-1}, and H was a submatrix of A.

Proofs of this fact generally use the Nullity Theorem, so we need to reconsider that theorem when $n = \infty$.

Nullity Theorem. Complementary submatrices of A and A^{-1} have the same nullity (dimension of nullspace).

If I and J are subsets of $1, \ldots, n$, then $A(I,J)$ is the submatrix containing the entries A_{ij} for i in I and j in J. The complementary submatrix contains the entries $(A^{-1})_{ij}$ for i not in J and j not in I.

When a is the upper left submatrix using the first i rows and j columns of A, its complementary submatrix Ca uses the last $n - j$ rows and $n - i$ columns of A^{-1}. When b is a lower left submatrix, Cb is also lower left (but its shape is transposed):

$$A = \begin{matrix} \begin{bmatrix} a & * \\ b & * \end{bmatrix} & \begin{matrix} i \\ n-i \end{matrix} \\ \begin{matrix} j & n-j \end{matrix} & \end{matrix} \qquad A^{-1} = \begin{matrix} \begin{bmatrix} * & * \\ Cb & Ca \end{bmatrix} & \begin{matrix} n-j \\ j \end{matrix} \\ \begin{matrix} n-i & i \end{matrix} & \end{matrix}.$$

The Nullity Theorem states that $\mathrm{nullity}(a) = \mathrm{nullity}(Ca)$ and $\mathrm{nullity}(b) = \mathrm{nullity}(Cb)$. There is an equivalent statement for the ranks, but that involves the size n of A and we want to allow $n = \infty$.

The history of this basic theorem is astonishingly recent. Our expository paper (8) attributed it to Gustafson in 1984. Now we find the equivalent theorem published by Kolotilina in the same year. Fiedler–Markham provided a matrix proof, and our favorite comes from Woerdeman (9). Horn and Johnson (10) give the Nullity Theorem an early place in their next edition. All these proofs start from block multiplication in $AA^{-1} = I = A^{-1}A$.

Kolotilina's second proof with Yeremin (11) is a new favorite. It begins with permutation matrices, when P^{-1} is simply P^{T}.

Nullity Theorem for Permutations. The nullity of an upper left submatrix p in P equals the nullity of the complementary lower right submatrix Cp in $P^{-1} = P^{\mathrm{T}}$.

Proof: If the upper left i by j submatrix has rank r, the nullity is $j - r$. Every row and column of P contains a single nonzero, so all ranks and nullities come from counting those 1s:

$$P = \begin{bmatrix} p & * \\ b & * \end{bmatrix} \quad \text{has nullities} \quad \begin{bmatrix} j-r & n-i-j+r \\ r & i-r \end{bmatrix}$$

$$P^{-1} = \begin{bmatrix} p^{\mathrm{T}} & b^{\mathrm{T}} \\ Cb & Cp \end{bmatrix} \quad \text{has nullities} \quad \begin{bmatrix} i-r & n-i-j+r \\ r & j-r \end{bmatrix}.$$

The nullities of p and Cp have the same value $j - r$, and the nullities of b and Cb have the same value r. The Nullity Theorem is proved for P by moving any submatrix $P(I,J)$ into a corner.

Kolotilina's insight was that the "Bruhat factorization" $A = LPU$ immediately gives the Nullity Theorem for A. Here U is upper triangular and L is lower triangular (we revise the standard Bruhat form by starting elimination at row 1). The key point is that P is in the middle (12, 13), unlike the usual factorization $PA = LU$ in numerical linear algebra. Then the upper left corner of $A = LPU$ is exactly $a = lpu$:

$$A = \begin{bmatrix} a & * \\ * & * \end{bmatrix} = \begin{bmatrix} l & 0 \\ * & * \end{bmatrix} \begin{bmatrix} p & * \\ * & * \end{bmatrix} \begin{bmatrix} u & * \\ 0 & * \end{bmatrix}.$$

Because l and u are invertible, $a = lpu$ has the same rank (and same nullity) as p.

The twin lemma for $A^{-1} = U^{-1}P^{-1}L^{-1}$ says that $Ca = (Cu)(Cp)(Cl)$. So Ca has the same nullity as Cp, which is the same as for p and for a.

Our point is simply that all steps of the proof remain valid for banded P and A and A^{-1} even when the matrices are infinite. We avoid doubly infinite matrices like the shift with $P_{i,i-1} = 1$ for $-\infty < i < \infty$, which could not be factored into a finite product of Fs. An alternative proof of our main theorem factors L, P, and U separately.

Summary

The two most active fields in applied linear algebra involve structured matrices and data matrices. Operators that arise in applications almost always have a special form—Toeplitz, Hankel, Laplacian, circulant, symplectic, …, Vandermonde, Hessenberg. Good algorithms (fast and stable) use those special structures. At the other extreme, huge matrices come from the floods of output in medical imaging and genomics and sensing of all kinds. There the goal is to find structure where none is apparent.

This paper and those to follow contribute to the analysis of one particular structure (perhaps the simplest): banded matrices. In this case ordinary elimination requires only w^2n steps, a crucial reduction from the familiar count $n^3/3$ for a full matrix. This linearity in n (sometimes $n \log n$) is typical of algorithms for structured matrices, and here it is easily recognized: w row operations act on rows of length w to eliminate one unknown at a time (all perfectly expressed by $A = LU$). Our count in the theorem above involved w^3 because the theorem allowed only 2 by 2 blocks, operating on adjacent rows.

In a future paper, one more group structure will be considered. The matrices $B = A + UV^{\mathrm{T}}$ are banded plus finite rank, with the requirement that A^{-1} is also banded. The Woodbury–Morrison formula expresses B^{-1} as A^{-1} plus finite rank, so we still have a group. This family is touching on the "semiseparable" and "quasiseparable" matrices that are now intensely studied.

Where banded matrices are the extreme case of rapid decay away from the diagonal, finite rank is the extreme case of an integral operator that is slowly varying. So we come closer to discrete forms of differential and integral equations. Here the model is Laplace's equation. If any single structured matrix can be identified as all-important in this corner of applied mathematics (perhaps small but astonishingly widespread), it is the graph Laplacian.

ACKNOWLEDGMENTS. We thank Vadim Olshevsky and Pavel Zhlobich for guiding us as they discovered "Green's matrices." We also thank Ilya Spitkovsky for simplifying the linear factors of $M(z)$ that succeed in construction 1 and now extend to construction 2.

1. Daubechies I (1988) Orthonormal bases of compactly supported wavelets. *Commun Pure Appl Math* 41:909–996.
2. Strang G, Nguyen T (1996) *Wavelets and Filter Banks* (Wellesley-Cambridge Press, Wellesley, MA).
3. Cantero MJ, Moral L, Velázquez L (2003) Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle. *Linear Algebra Appl* 362:29–56.
4. Gohberg I, Kaashoek MA, Spitkovsky I (2003) An overview of matrix factorization theory and operator applications: Factorization and integrable systems. *Oper Th Adv Appl* 141 (Birkhäuser, Basel), 1–102.
5. Olshevsky V, Zhlobich P, Strang G (2010) Green's matrices. *Linear Algebra Appl* 432:218–241.
6. Kimura H (1985) Generalized Schwarz form and lattice-ladder realization of digital filters. *IEEE Trans Circuits Syst* 32:1130–1139.
7. Asplund E (1959) Inverses of matrices $\{a_{ij}\}$ which satisfy $a_{ij} = 0$ for $j > i + p$. *Math Scand* 7:57–60.
8. Strang G, Nguyen T (2004) The interplay of ranks of submatrices. *SIAM Rev* 46:637–646.
9. Woerdeman H (2008) A matrix and its inverse: Revisiting minimal rank completions, recent advances in matrix and operator theory. *Oper Th Adv Appl* 179:329–338.
10. Horn R, Johnson C (1985) *Matrix Analysis* (Cambridge University Press, Cambridge, UK).
11. Kolotilina LY, Yeremin AY (1987) Bruhat decomposition and solution of sparse linear algebra systems. *Sov J Numer Anal Math Modelling* 2:421–436.
12. Elsner L (1979) On some algebraic problems in connection with general eigenvalue algorithms. *Linear Algebra Appl* 26:123–138.
13. Gohberg I, Goldberg S (1982) Finite dimensional Wiener–Hopf equations and factorizations of matrices. *Linear Algebra Appl* 48:219–236.