CONVERGENCE OF LR ALGORITHM FOR A ONE-POINT SPECTRUM TRIDIAGONAL MATRIX

CARLA FERREIRA* AND BERESFORD PARLETT†

Abstract. We prove convergence for the basic LR algorithm on a real unreduced tridiagonal matrix with a one-point spectrum - the Jordan form is one big Jordan block. First we develop properties of eigenvector matrices. We also show how to deal with the singular case.

Key words. unsymmetric tridiagonal matrices, multiple eigenvalues, LR algorithm

AMS subject classifications. 65F15

1. Introduction.

This paper presents a rigorous proof that the LR algorithm, without shifts, applied to an unreduced tridiagonal matrix with a one-point spectrum converges to an upper bidiagonal matrix. The rate of convergence is very slow, like $1/k$ after k steps, but what is remarkable is that the algorithm actually converges. We hasten to say that this result is not exactly new. In the middle of the 1960's J.H. Wilkinson sketched out the underlying reason for this surprising result, both for LR and QR, but he was not concerned with tridiagonal matrices and he needed assumptions that the column and row eigenvector matrices were completely regular. Moreover, he did not show that a certain universal matrix was also completely regular.

So the contribution of this paper is twofold. We show that in the unreduced tridiagonal case the eigenvector matrices are completely regular and we show that the universal matrix mentioned above is also completely regular, not just in the asymptotic regime. In contrast to most papers, the focus is not on the result but on the proof.

The reason for considering the LR algorithm instead of the more popular QR is that it preserves tridiagonal form. The fear of instability which undermined the adoption of the LR algorithm is not justified. In the tridiagonal case the new iterate need not overwrite the old one; instead the new one can be stored separately and, if element growth is unacceptable, then it is rejected, the shift is modified (usually reduced) and the transform is reapplied. A reward for this approach is that it encourages more aggressive and powerful shift strategies than were used in the past. However, implementation details are not part of this paper.

For the sake of brevity this paper is addressed to readers who are already familiar with the LR and QR algorithms including their convergence properties when the eigenvalues have distinct moduli. See [1, 4, 12] for such material. Some readers may enjoy the detailed example of a 6×6 tridiagonal with a one-point spectrum and the choice of generalized eigenvectors.

In the absence of breakdown, the basic LR algorithm is given by

$A_1 = A$
for $i = 1, 2, \ldots$
  Factor $A_i = L_i R_i$ (L_i unit lower triangular, U_i upper triangular)
  $A_{i+1} = R_i L_i$
end

*Mathematics Department, University of Minho, 4710-057 Braga (caferrei@math.uminho.pt)

†Department of Mathematics and the Computer Science Division of the Electrical Engineering and Computer Science Department, University of California, Berkeley, California 94720 (parlett@math.berkeley.edu).

We recall two key facts: for $i = 1, 2, \dots,$

$$A_{i+1} = {\mathcal{L}_i}^{-1} A \mathcal{L}_i \qquad \text{with} \qquad \mathcal{L}_i \equiv L_1 L_2 \dots L_i, \tag{1.1}$$

and

$$A^i = \mathcal{L}_i \mathcal{U}_i \qquad \text{with} \qquad \mathcal{U}_i \equiv R_i R_{i-1} \dots R_1. \tag{1.2}$$

2. Eigenvector properties of a one-point spectrum tridiagonal. When λ is a multiple eigenvalue the eigenvector matrix must be filled out with the so-called generalized eigenvectors with the property that, for any such C,

$$(C - \lambda I)^j v = 0, \;\; (C - \lambda I)^{j-1} v \neq 0.$$

We say that v is an eigenvector of grade j and omit the word generalized in the rest of this paper.

In what follows we shall present some properties of eigenvector matrices that are sufficient to guarantee convergence of the basic LR algorithm without invoking the extra hypotheses needed by Rutishauser [11] and Wilkinson [13, 14, pp.487-492] for the general case. To the best of our knowledge these results are new.

Following standard usage in Linear System Theory we say that X is *completely* (or *strongly*) *regular* when X and all its leading principal submatrices are nonsingular. We shall use the terms "completely regular" and "permits LU" interchangeably. To be precise, we note that a singular matrix may permit triangular factorization but in our work all the matrices of interest will be nonsingular.

Most of our results extend directly to complex unreduced tridiagonal matrices but we focus on real matrices for simplicity and because it is the most frequent case in applications.

Consider a real tridiagonal matrix

$$C = \begin{bmatrix} a_1 & c_1 & & & \\ b_1 & a_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & b_{n-2} & a_{n-1} & c_{n-1} \\ & & & b_{n-1} & a_n \end{bmatrix} \in \mathbb{R}^{n \times n} \tag{2.1}$$

with $b_i c_i \neq 0$, $i = 1, \dots, n-1$. With this constraint we say that C is *unreduced.*

Define monic polynomials $p_0, p_1, \dots, p_n$ by

$$p_0(\tau) = 1, \;\; p_j(\tau) := \det(\tau I_j - C_j), \;\; j = 1, \dots, n,$$

where I_j represents the $j \times j$ identity matrix and C_j the j^{th} leading principal submatrix of C.

The celebrated three term recurrence (3TR) for C is

$$\begin{aligned} p_1(\tau) &= (\tau - a_1) = (\tau - a_1) p_0(\tau), \\ p_{j+1}(\tau) &= (\tau - a_{j+1}) p_j(\tau) - b_j c_j p_{j-1}(\tau), \qquad j = 1, 2, \dots, n-1. \end{aligned}$$

In this paper we suppose that C's spectrum consists of a single nonzero point λ and that its Jordan form is

$$J = \lambda I + N$$

where N is the nilpotent matrix

$$N = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}.$$

Definition 2.1.

$$\boldsymbol{p}(\lambda) := \begin{bmatrix} p_0(\lambda) & p_1(\lambda) & \dots & p_{n-1}(\lambda) \end{bmatrix}^T, \qquad j = 1, \dots, n.$$

Definition 2.2. *We will denote*

$$D_b := diag(1, b_1, b_1b_2, b_1b_2b_3, \dots, b_1b_2 \cdots b_{n-1}), \tag{2.2}$$
$$D_c := diag(1, c_1, c_1c_2, c_1c_2c_3, \dots, c_1c_2 \cdots c_{n-1}). \tag{2.3}$$

The 3TR is equivalent to the matrix equation

$$\left(D_c C D_c^{-1} - \tau I\right) \boldsymbol{p}(\tau) = -\boldsymbol{e}_n p_n(\tau).$$

Pre-multiplying both sides by D_c^{-1} we get

$$\left(C - \tau I\right) D_c^{-1} \boldsymbol{p}(\tau) = -\boldsymbol{e}_n \frac{p_n(\tau)}{c_1 \dots c_{n-1}}. \tag{2.4}$$

When $\tau = \lambda$,

$$(C - \lambda I) D_c^{-1} \boldsymbol{p}(\lambda) = \boldsymbol{0}$$

and we see that the only column eigenvector of C is $D_c^{-1}\boldsymbol{p}(\lambda)$. Similarly, its single row eigenvector is $\boldsymbol{p}(\lambda)^T D_b^{-1}$.

One way to find eigenvectors of higher grade is to differentiate (2.4) as many times as is necessary. Differentiate once to get

$$\left(C - \tau I\right) D_c^{-1} \boldsymbol{p}'(\tau) - D_c^{-1} \boldsymbol{p}(\tau) = -\boldsymbol{e}_n \frac{p_n'(\tau)}{c_1 \dots c_{n-1}}.$$

After taking k derivatives we have

$$\left(C - \tau I\right) D_c^{-1} \boldsymbol{p}^{(k)}(\tau) - k D_c^{-1} \boldsymbol{p}^{(k-1)}(\tau) = -\boldsymbol{e}_n \frac{p_n^{(k)}(\tau)}{c_1 \dots c_{n-1}}. \tag{2.5}$$

Dividing through by $k!$ we obtain

$$\left(C - \lambda I\right) D_c^{-1} \frac{\boldsymbol{p}^{(k)}(\lambda)}{k!} = D_c^{-1} \frac{\boldsymbol{p}^{(k-1)}(\lambda)}{(k-1)!}, \qquad k = 0, 1, \dots, n-1. \tag{2.6}$$

This is valid for any unreduced tridiagonal C.

The unit lower triangular matrix

$$P = \left[\boldsymbol{p}(\lambda) \quad \boldsymbol{p}'(\lambda) \quad \tfrac{1}{2!}\boldsymbol{p}''(\lambda) \quad \dots \quad \tfrac{1}{(n-1)!}\boldsymbol{p}^{(n-1)}(\lambda)\right] \tag{2.7}$$

plays an important role in our analysis. P and P^T are called *polynomial Vandermonde* matrices. Now we have one possible eigenvector matrix,

$$CD_c^{-1}P = D_c^{-1}P(N + \lambda I). \tag{2.8}$$

A 6×6 example of P is given in the following section 2.1. A corresponding matrix of row eigenvectors is *not* $P^T D_b^{-1}$.

To find the row eigenvectors for C introduce the notation

$$\not{I} = \begin{bmatrix} & & & & 1 \\ & & & 1 & \\ & & \iddots & & \\ & 1 & & & \\ 1 & & & & \end{bmatrix}.$$

For C^T (2.8) is

$$C^T D_b^{-1} P = D_b^{-1} P(N + \lambda I).$$

Transpose, replace N^T by $\not{I} N \not{I}$ and pre-multiply by $\not{I}$ to find

$$\left(\not{I} P^T D_b^{-1}\right) C = (N + \lambda I)\left(\not{I} P^T D_b^{-1}\right). \tag{2.9}$$

So, $\not{I} P^T D_b^{-1}$ is the matrix of row eigenvectors of C.

Recall that $D_c^{-1}P$ is lower triangular and $P^T D_b^{-1}$ is upper triangular. Nevertheless, it is not true that the product $\left(\not{I} P^T D_b^{-1}\right)\left(D_c^{-1}P\right)$ is diagonal. The reason is subtle: for a Jordan block, the eigenvectors of grade higher than 1 are not uniquely defined. One may add to an eigenvector of grade k any multiple of any eigenvector of lower grade. In matrix terms, we may post-multiply $D_c^{-1}P$ by any unit upper triangular matrix $\mathcal{U}$. However to preserve the 1's in the Jordan form, $\mathcal{U}$ has to be Toeplitz. All suitable matrices are of the form $\varphi(N)$, φ a polynomial with degree $< n$, which satisfy $\varphi(O) = I$. Thus, $\varphi(N)$ commutes with $\lambda I + N$ and

$$CD_c^{-1}P\varphi(N) = D_c^{-1}P(\lambda I + N)\varphi(N) = D_c^{-1}P\varphi(N)(\lambda I + N).$$

That is, $D_c^{-1}P$ is only unique up to post-multiplication by a nonsingular polynomial $\varphi(N)$. The preferred choice of $\varphi(N)$ for us is given by

$$\left(\not{I} P^T D_b^{-1}\right)\left(D_c^{-1}P\right) = \varphi(N). \tag{2.10}$$

We have prooved

Theorem 2.3. *If unreduced tridiagonal C has one-point spectrum λ and P, D_b, D_c are as defined in (2.7), (2.2) and (2.3) then*

$$C = D_c^{-1}P\varphi(N)^{-1}(\lambda I + N)\not{I} P^T D_b^{-1},$$

for a certain polynomial φ with $\varphi(O) = I$, determined by (2.10).

The following property of $\varphi(N)$ will be needed later.

LEMMA 2.4. *The matrix $\varphi(N)^{-1}\not{I}$ admits triangular factorization:*

$$\varphi(N)^{-1}\not{I} = \left(P^{-1}\right)\left(D_c D_b\right)\left(P^{-T}\right).$$

Proof. Invert (2.10). □

We found the example that follows helpful in understanding the role of $\varphi(N)$ in this theorem.

2.1. Example of a one-point spectrum tridiagonal. Recall that a square matrix A is *Toeplitz* when the entries of A are constant down the diagonals parallel to the main diagonal and is *Hankel* when the entries of A are constant along the diagonals perpendicular to the main diagonal.

In [6] Z. S. Liu devised an algorithm to obtain unreduced tridiagonal matrices with one-point spectrum of arbitrary dimension $n \times n$. These matrices, that we will call *Liu matrices*, have only one eigenvalue, zero, with algebraic multiplicity n and geometric multiplicity 1. The Jordan form consists of one big Jordan block. We will represent Liu matrices as

$$Liu_n = \text{tridiag}(\mathbf{1}^n, \boldsymbol{\alpha}^n, \boldsymbol{\gamma}^n)$$

where $\mathbf{1}^n$ always stands for a vector of 1's of length $n-1$. For $n = 6$, $\boldsymbol{\alpha}^6 = \begin{bmatrix}0 & 0 & -1 & 1 & 0 & 0\end{bmatrix}$ and $\boldsymbol{\gamma}^6 = \begin{bmatrix}-1 & 1 & -1 & 1 & -1\end{bmatrix}$.

The transpose is more convenient,

$$Liu_6^T = \begin{bmatrix} 0 & 1 & & & & \\ -1 & 0 & 1 & & & \\ & 1 & -1 & 1 & & \\ & & -1 & 1 & 1 & \\ & & & 1 & 0 & 1 \\ & & & & -1 & 0 \end{bmatrix}.$$

We have

$$\begin{aligned}
p_0(\tau) &= 1, \\
p_1(\tau) &= \tau, \\
p_2(\tau) &= \tau^2 + 1, \\
p_3(\tau) &= (\tau + 1)p_2(\tau) - p_1(\tau) = \tau^3 + \tau^2 + 1, \\
p_4(\tau) &= (\tau - 1)p_3(\tau) + p_2(\tau) = \tau^4 + \tau, \\
p_5(\tau) &= \tau p_4(\tau) - p_3(\tau) = \tau^5 - \tau^3 - 1, \\
p_6(\tau) &= \tau p_5(\tau) + p_4(\tau) = \tau^6.
\end{aligned}$$

Then

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & -1 & 0 & 1 \end{bmatrix}, \qquad D_b = \text{diag}(1, -1, -1, 1, 1, -1), \qquad D_c = I.$$

Now we can define $\mathcal{U}$ by

$$P^T D_b^{-1} D_c^{-1} P = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & -1 \end{bmatrix} = \text{Ɨ}\mathcal{U}.$$

Thus,

$$\mathcal{U} = \text{Ɨ} P^T D_b^{-1} D_c^{-1} P = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathcal{U}^{-1} = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

We see that $\mathcal{U}$ is unit upper triangular and Toeplitz: $\mathcal{U} = I + N^3 - N^5 = \varphi(N)$, a polynomial in N that commutes with $\lambda I + N$ (also $\mathcal{U}^{-1} = I - N^3 + N^5$). Finally, the spectral decomposition of Liu_6^T:

$$Liu_6^T = D_c^{-1} P \mathcal{U}^{-1} \,(0I + N)\, \text{Ɨ} P^T D_b^{-1}, \qquad I = \left(\text{Ɨ} P^T D_b^{-1}\right)(D_c^{-1} P \mathcal{U}^{-1}).$$

Note that $P\mathcal{U}^{-1}$ is in LU form and $P^T D_b^{-1}$ is upper triangular.

3. Convergence of basic LR algorithm on a one-point spectrum tridiagonal. How can the analysis of the distinct absolute value case (see Wilkinson [13] and [14, pp. 487-492]) be rebuilt when an eigenvalue is multiple so that no shift will produce different moduli? We will deal first with the case $\lambda \neq 0$.

3.1. The case $\boldsymbol{\lambda \neq 0}$. The unit lower triangular L factor of a matrix M will be denoted by $\mathcal{L}(M)$ and the upper triangular U factor by $\mathcal{U}(M)$, when M is completely regular. In this notation

$$C^k = \mathcal{L}(C^k)\mathcal{U}(C^k) \qquad \text{and} \qquad C_{k+1} = \mathcal{L}(C^k)^{-1} C \mathcal{L}(C^k) \tag{3.1}$$

where C_{k+1} is the LR transform after k steps.

Recall that the Vandermonde matrix P for the one-point spectrum case is unit lower triangular and from (2.8) and theorem 2.3 see that

$$C = X(\lambda I + N)X^{-1} \tag{3.2}$$

with

$$X = D_c^{-1} P \varphi(N)^{-1} \qquad \text{and} \qquad X^{-1} = \text{Ɨ} P^T D_b^{-1}$$

where $\varphi(N)$ is given in (2.10). Then

$$C^k = D_c^{-1} P \varphi(N)^{-1} (\lambda I + N)^k \, \text{Ɨ} P^T D_b^{-1}. \tag{3.3}$$

Note that P^T is unit upper triangular.

Our method of proof is in the same spirit as sketches in Wilkinson's book [14, pp. 517-519 and 521-522] for the general case. He was not concerned with tridiagonal matrices and had to assume explicitly that the column and row eigenvector matrices X and X^{-1} of C were completely regular. These assumptions are no longer needed for an unreduced tridiagonal matrix. The following lemma is the key for the algorithm not to fail.

Lemma 3.1. *For all $k \geq n$, $(\lambda I + N)^k \not{I}$ for $\lambda \neq 0$ is completely regular and thus admits triangular factorization, say*

$$(\lambda I + N)^k \not{I} = L_k D_k \lambda^k L_k^T,$$

and, as $k \to \infty$, $L_k = I + E_k$, $E_k \to O$. The rate of convergence is low, $\mathcal{O}(1/k)$.

The proofs of this lemma and Theorem 3.3, from which the lemma follows, will be given in the next section.

Theorem 3.2. *Let C be a nonsingular unreduced tridiagonal matrix that permits triangular factorization and has a one-point spectrum $\lambda \neq 0$. Given the notation above, the basic* LR *algorithm applied to C produces a sequence of matrices C_k that converges (in exact arithmetic) to*

$$D_c^{-1}(\lambda I + N)D_c$$

with D_c defined in (2.3).

Proof. The proof manipulates C^k into LU form. For $k \geq n$, insert Lemma 3.1's result into (3.3) to get

$$\begin{aligned} C^k &= D_c^{-1} P \left(D_c D_c^{-1}\right) \varphi(N)^{-1}(I + E_k) D_k \lambda^k L_k^T P^T D_b^{-1} \\ &= D_c^{-1} P D_c (I + F_k) D_c^{-1} \varphi(N)^{-1} D_k \lambda^k L_k^T P^T D_b^{-1} \end{aligned}$$

with

$$F_k = \left(\varphi(N) D_c\right)^{-1} E_k \left(\varphi(N) D_c\right) \to O \qquad \text{as} \qquad k \to \infty,$$

since

$$\|F_k\| \leq \operatorname{cond}\left(\varphi(N) D_c\right) \|E_k\| \to 0 \qquad \text{as} \qquad k \to \infty.$$

Thus,

$$\mathcal{L}(C^k) = D_c^{-1} P D_c \mathcal{L}(I + F_k) \to D_c^{-1} P D_c \qquad \text{as} \qquad k \to \infty,$$

since $D_c^{-1}\varphi(N)^{-1} D_k \lambda^k L_k^T P^T D_b^{-1}$ is upper triangular ($\varphi(N)^{-1}$ is upper triangular and Toeplitz).

Finally, since P is unit lower triangular, the LU factorization of $X = D_c^{-1} P \varphi(N)^{-1}$ is

$$X = \left(D_c^{-1} P D_c\right)\left(D_c^{-1} \varphi(N)^{-1}\right) =: L_X U_X$$

and then

$$\begin{aligned} C_{k+1} &= \mathcal{L}(C^k)^{-1} C \mathcal{L}(C^k) \\ &= \mathcal{L}(C^k)^{-1} X (\lambda I + N) X^{-1} \mathcal{L}(C^k) && \text{(by 3.2)} \\ &\quad \to \left(D_c^{-1} P D_c\right)^{-1} X (\lambda I + N) X^{-1} \left(D_c^{-1} P D_c\right) \\ &= L_X^{-1} L_X U_X (\lambda I + N) U_X^{-1} L_X^{-1} L_X \\ &= U_X (\lambda I + N) U_X^{-1} \\ &= D_c (\lambda I + N) D_c^{-1}, \end{aligned}$$

since $\varphi(N)$ commutes with $\lambda I + N$. Notice that $D_c(\lambda I+N)D_c^{-1}$ is *not* Toeplitz unless $D_c = I$. It is satisfying that $\varphi(N)$ cancels and does not influence the limit. □

Can the LR algorithm applied to our special matrices breakdown in the early stages? The answer is yes, but only if $|\lambda|$ is small. This feature depends on the property of $\varphi(N)\not{I}$ expressed in Lemma 2.4. Consider, for low values of k,

$$C^k = D_c^{-1}P(\lambda I+N)^k\varphi(N)^{-1}\not{I}P^TD_b^{-1}.$$

We want to know whether $(\lambda I+N)^k\varphi(N)^{-1}\not{I}$ permits triangular factorization. Observe that $(\lambda I+N)^k$ is a polynomial in λ with leading term $I\lambda^k$. Thus the leading term in $(\lambda I+N)^k\varphi(N)^{-1}\not{I}$ is $\varphi(N)^{-1}\not{I}\lambda^k$ for $k = 1, 2, 3, \dots$. For large enough $|\lambda|$ this term dominates the rest and so, by Lemma 2.4, $(\lambda I+N)^k\varphi(N)^{-1}\not{I}$ permits triangular factorization and thus LR algorithm is well defined. A few experiments suggest that LR does not breakdown (in our case) when $|\lambda| > 1$.

3.1.1. Proof of Lemma 3.1. Since $N^i = O$ for $i \geq n$, $(\lambda I+N)^k$ is a polynomial of degree $n-1$, for $k \geq n$, with coefficients that depend on k. From [5, p.138]

$$(\lambda I+N)^k = \sum_{i=0}^{n-1}\binom{k}{i}\lambda^{k-i}N^i = \lambda^k\Delta_\lambda\left(\sum_{i=0}^{n-1}\binom{k}{i}N^i\right)\Delta_\lambda^{-1}$$

where Δ_λ is the matrix defined as

$$\Delta_\lambda := \mathrm{diag}(1, \lambda, \lambda^2, \dots, \lambda^{n-1}).$$

So $(\lambda I+N)^k\not{I}$ is Hankel and upper anti-triangular. Define

$$H_k = \left(\sum_{i=0}^{n-1}\binom{k}{i}N^i\right)\not{I} = \begin{bmatrix} \binom{k}{n-1} & \binom{k}{n-2} & \binom{k}{n-3} & \dots & \binom{k}{2} & k & 1 \\ \binom{k}{n-2} & \binom{k}{n-3} & \dots & \binom{k}{2} & k & 1 & 0 \\ \binom{k}{n-3} & \dots & \binom{k}{2} & k & 1 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \binom{k}{2} & k & 1 & 0 & 0 & \dots & 0 \\ k & 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 & 0 \end{bmatrix}$$

so that

$$(\lambda I+N)^k\not{I} = \Delta_\lambda H_k\not{I}\Delta_\lambda^{-1}\not{I}\lambda^k.$$

By Theorem 3.3 and Corollary 3.5 below, H_k is completely regular and admits the factorization

$$H_k = \widetilde{L}_k\widetilde{D}_k\widetilde{L}_k^T \qquad \text{with} \qquad \widetilde{L}_k = I + \widetilde{E}_k, \quad \widetilde{E}_k \to O \text{ as } k \to \infty.$$

Thus, for each $k \geq n$,

$$(\lambda I+N)^k\not{I} = L_kD_k\lambda^kL_k^T \qquad \text{with} \qquad L_k = I + \Delta_\lambda\widetilde{E}_k\Delta_\lambda^{-1} = I + E_k \to I.$$

The diagonal matrix $D_k\lambda^k$ is also a function of k but it may not converge to a finite matrix. □

The formulae that follow are the outcome of a difficult determinantal evaluation and we have not found them in the literature. Unfortunately, we were not able to show that H_k is completely regular without exhibiting the actual formulae for each determinant.

For $1 \le p \le n$ let $H_n(k)_p$ designate the leading principal $p \times p$ submatrix of $n \times n$ H_k. Also we define the double factorial as follows:

$$m!! = m!(m-1)!(m-2)!\cdots 2!1!, \qquad m \in \mathbb{N}$$
$$0!! = 1.$$

THEOREM 3.3. *Let $n \in \mathbb{N}$, $1 \le p \le n$ and $l = \min(p, n-p)$. Define*

$$c_{n,l} = \begin{cases} \prod_{j=1}^{\frac{1}{2}l-1} \left[(n-2j)(n-2j-1)\right]^j \cdot m_l \cdot \prod_{j=\frac{1}{2}l+1}^{l-1} \left[(n-2j+1)(n-2j)\right]^{l-j} /(l-1)!! \\ \quad \text{with } m_l = (n-l)^{\frac{1}{2}l}, \qquad \text{if } l \text{ is even} \\ \prod_{j=1}^{\frac{1}{2}(l-3)} \left[(n-2j)(n-2j-1)\right]^j \cdot m_l \cdot \prod_{j=\frac{1}{2}(l+3)}^{l-1} \left[(n-2j+1)(n-2j)\right]^{l-j} /(l-1)!! \\ \quad \text{with } m_l = \left[(n-l+1)(n-l)(n-l-1)\right]^{\frac{1}{2}(l-1)}, \qquad \text{if } l \text{ is odd} \end{cases}$$

Then

$$\det\left(H_n(k)_p\right) = \frac{s}{c_{n,l}} \prod_{i=1}^{l} \binom{k+p-i}{n-2i+1}, \qquad 1 \le p < n,$$
$$\det\left(H_n(k)_n\right) = s$$

where

$$s = \begin{cases} 1 & \text{if } p \equiv 0,1 \pmod 4 \\ -1 & \text{if } p \equiv 2,3 \pmod 4 \end{cases}.$$

Once discovered, these formulae have been verified using MATHEMATICA for various values of n.

As an example, if $n = 8$ the determinants $\det\left(H_n(k)_p\right)$ for $p = 1 : n-1$ are

$$\binom{k}{n-1}, \quad -\frac{1}{6}\binom{k+1}{n-1}\binom{k}{n-3}, \quad -\frac{1}{60}\binom{k+2}{n-1}\binom{k+1}{n-3}\binom{k}{n-5},$$
$$\frac{1}{240}\binom{k+3}{n-1}\binom{k+2}{n-3}\binom{k+1}{n-5}\binom{k}{n-7},$$
$$\frac{1}{60}\binom{k+4}{n-1}\binom{k+3}{n-3}\binom{k+2}{n-5}, \quad -\frac{1}{6}\binom{k+5}{n-1}\binom{k+4}{n-3}, \quad -\binom{k+6}{n-1}$$

We see that the $c_{n,l}$ are

$$1, \quad 6, \quad 60, \quad 240, \quad 60, \quad 6, \quad 1.$$

Remarks on the proof of Theorem 3.3. We sketch our method for evaluating the determinants of $H_n(k)_p$. The first step is the unobvious one; it destroys the Hankel form in a useful way. It employs over and over again the basic identity

$$\binom{a}{b-1} + \binom{a}{b} = \binom{a+1}{b}.$$

We treat the case $p < n/2$.

Step 1. For $j = 1, 2, \ldots, p-1$ add column $j+1$ to column j. The result is that all k's in columns $1 : p-1$ became $k+1$. Next, for $j = 1, 2, \ldots, p-2$ add column $j+1$ to column j so that each $k+1$ in columns $1 : p-2$ became $k+2$. Continue this process to obtain a new matrix with the same determinant,

$$\begin{bmatrix} \binom{k+p-1}{n-1} & \binom{k+p-2}{n-2} & \cdots & \binom{k+1}{n-p+1} & \binom{k}{n-p} \\ \binom{k+p-1}{n-2} & \binom{k+p-2}{n-3} & \cdots & \binom{k+1}{n-p} & \binom{k}{n-p-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \binom{k+p-1}{n-p} & \binom{k+p-2}{n-p-1} & \cdots & \binom{k+1}{n-2p+2} & \binom{k}{n-2p+1} \end{bmatrix}.$$

Step 2. Expressing the binomial coefficients as factorials we can remove all factors involving k from rows and columns. In addition, binomial coefficients can be recovered by factoring $1/(n-2j+1)!$ from each row. This yields the common factor

$$\prod_{j=1}^{p} \binom{k+p-j}{n-2j+1}$$

and leaves a strange matrix $K_p^{(n)}$ whose (i,j) entry is $(n-2i+1)!/(n-2j+1)!$. It turns out that

$$\det\left(K_p^{(n)}\right) = s/c_{n,l}$$

where $l = \min(p, n-p)$. □

Lemma 3.4. *For $1 \le p \le n$ and $k \ge n$, $H_n(k)_p$ is completely regular.*

Proof. For $k \ge n$, $k+p-i \ge n-2i+1 > 0$ and so all the binomial coefficients are positive. □

Corollary 3.5. *H_k admits triangular factorization*

$$H_k = \widetilde{L}_k \widetilde{D}_k \widetilde{L}_k^T.$$

The subdiagonal entries of $\widetilde{L}_k$ are given by

$$l_{j+1,j} = \frac{j(n-j)}{k-n+2j}, \qquad j = 1, 2, \ldots, n-1,$$

and, as $k \to \infty$,

$$l_{j+m,j} = \mathcal{O}\left(k^{-m}\right), \qquad m = 2, 3, \ldots, n-1; \;\; j = 1, 2, \ldots, n-m.$$

Proof. Cramer's rule shows $l_{j+m,j}$ as a quotient of monic polynomials in k whose degrees differ by m. □

As an example, for $n = 6$, the $l_{j+1,j}$ entries of $\widetilde{L}_k$ are

$$\frac{5}{k-4}, \quad \frac{8}{k-2}, \quad \frac{9}{k}, \quad \frac{8}{2+k}, \quad \frac{5}{4+k}.$$

So, for $m \geq 1$, the $(j+m, j)$ entry of $\widetilde{L}_k$ is $\mathcal{O}\big(k^{-m}\big)$ and thus

$$\widetilde{L}_k \to I + \widetilde{E}_k, \quad \widetilde{E}_k \to O \qquad \text{as} \qquad k \to \infty.$$

But the convergence is very slow, governed by $\mathcal{O}\left(k^{-1}\right)$.

3.2. The case $\boldsymbol{\lambda = 0}$. The matrix C is nilpotent so that C^k vanishes for $k \geq n$. Thus the LR algorithm is neither well defined nor needed. Nevertheless, this is an important case and must be examined. Only in 2008 did we realize that the algorithm below gives an ideal prologue to a tridiagonal eigensolver because it wastes a small amount of effort on standard cases and deals accurately and efficiently with difficult cases such as Liu matrices.

What happens if C does not permit triangular factorization and yet is singular? The solution is surprisingly simple. The long abandoned Givens' method for computing an eigenvector solves $Cx = 0$ by assuming $x_1 = 1$ and using row j to determine x_{j+1} for $j = 1, 2, \ldots, n-1$. The last equation

$$c_{n,n-1} x_{n-1} + c_{nn} x_n = 0$$

will be satisfied when and only when C is singular.

The next step is to set $x^{(1)} = x$ and try to solve $Cx^{(2)} = x^{(1)}$ with starting value $x_1^{(2)} = 0$. Thus, as before, use row j to determine $x_{j+1}^{(2)}$ for $j = 1, 2, \ldots, n-1$. If $\lambda = 0$ has multiplicity ≥ 2, then the last equation

$$c_{n,n-1} x_{n-1}^{(2)} + c_{nn} x_n^{(2)} = x_n^{(1)}$$

will be satisfied and the process continues:

for $k = 3, 4, \ldots$ **do**
  set
    $$x_1^{(k)} = x_2^{(k)} = \cdots = x_{k-1}^{(k)} = 0$$
  solve
    $$Cx^{(k)} = x^{(k-1)} \qquad \text{by Givens' method}$$
  until
    $$c_{n,n-1} x_{n-1}^{(k)} + c_{nn} x_n^{(k)} \neq x_n^{(k-1)} \quad \textbf{\textit{or}} \quad k = n+1$$

Upon exit, the multiplicity of $\lambda = 0$ is revealed as $k-1$ and

$$x^{(1)}, x^{(2)}, \ldots, x^{(k-1)}$$

form a Jordan chain for $\lambda = 0$.

This procedure suggests that if unreduced matrix C has a one-point spectrum then the eigenvalue mean

$$\overline{\lambda} = \operatorname{trace}(C)/n$$

will be the spectral point and the adaptation of Givens' method to $C - \bar{\lambda} I$ described above will yield a Jordan basis without the need for the LR algorithm.

However the LR algorithm is useful for the general case and our analysis shows that even without the optimal shift the algorithm converges for multiple eigenvalues.

REFERENCES

[1] J. W. Demmel, *Applied Numerical Linear Algebra*, Society for Industrial and Applied Mathematics, 1997.

[2] M. Fiedler, *Special matrices and their applications in numerical mathematics*, Martinus Nijhoff Publishers, 1986.

[3] F. R. Gantmacher and M. G. Krein, *Oscillation matrices and kernels and small vibrations of mechanical systems*, United States Atomic Energy Commission, Office of Technical Information. Translated from a publication of the State Publishing House for Technical-Theoretical Literature, Moscow-Leningrad, 1950.

[4] G. H. Golub and F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore and London, 3rd ed., 1996.

[5] R. A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge University Press, 1996.

[6] Z. S. Liu, *On the extended HR algorithm*, Technical Report PAM-564, Center for Pure and Applied Mathematics, University of California, Berkeley, CA, USA, 1992.

[7] B. N. Parlett, *The development and use of methods of LR type*, SIAM Review, 6:275-295, 1964.

[8] ———, *Canonical Decomposition of Hessenberg Matrices*, Mathematics of Computation, 98(vol.21):223-227, 1967.

[9] ———, *Global Convergence of the Basic QR algorithm on Hessenberg Matrices*, Mathematics of Computation, 104(vol.22):803-817, 1968.

[10] ———, *What Hadamard Missed*, Unpublished Technical Report, 1996.

[11] H. Rutishauser and H.R. Schwarz, *The LR transformation method for symmetric matrices*, Numerische Mathematic, 5:273-289, 1963.

[12] D. S. Watkins, *QR-like algorithms - an overview of convergence theory and practice*, Lectures in Applied Mathematics, 32:879-893, 1996.

[13] J. Wilkinson, *Convergence of the LR, QR, and related algorithms*. Computing Journal, 8:77-84, 1965.

[14] ———, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.

[15] ———, *Global convergence of tridiagonal QR algorithm with origin shifts*, Linear Algebra and Its Applications, 1:409-420, 1968.