CONVERGENCE OF LR 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 develope properties of eigenvector matrices.

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 undetermined 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.

The reader is supposed to have some prior exposure to the QR and/or LR algorithms. We omit proofs of well known results.

Some readers may enjoy the detailed example of a 6×6 tridiagonal with a one-point spectrum.

2. Classical results for the convergence of LR algorithm .

The k^{th} iteration of the basic LR method is based on the LU decomposition $A_k = L_k U_k$ and on the multiplication of the factors L_k and U_k in reverse order to get the matrix A_{k+1}. The sequence of matrices $A =: A_1, A_2, A_3, \ldots$ is then generated by

LR algorithm

$A_1 = A$
for $i = 1, 2, \ldots$
  Factor $A_i = L_i R_i$ (LU factorization)
  $A_{i+1} = R_i L_i$
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).

It is easy to see that A_{k+1} is similar to A_k,

$$A_{k+1} = R_k L_k = L_k^{-1}(L_k R_k) L_k,$$

and then, by an inductive argument, A_{k+1} is similar to A.

LEMMA 2.1. *Let $A_1 := A$ and $\{A_i\}_{i=1}^{\infty}$ be the sequence of matrices generated by LR algorithm. Then*

$$A_{i+1} = (L_1 L_2 \dots L_i)^{-1} A (L_1 L_2 \dots L_i), \qquad i = 1, 2, \dots \dots$$

LEMMA 2.2. *Let matrices $\mathcal{L}_i$ and $\mathcal{U}_i$ be defined by*

$$\mathcal{L}_i \equiv L_1 L_2 \dots L_i \qquad \text{and} \qquad \mathcal{U}_i \equiv R_i R_{i-1} \dots R_1.$$

Then $\mathcal{L}_i \mathcal{U}_i$ is the LU *decomposition of A^i.*

Notice the reverse ordering in $\mathcal{L}_i$ and $\mathcal{U}_i$.

Thus, we have

$$A_{i+1} = \mathcal{L}_i^{-1} A \mathcal{L}_i \qquad \text{and} \qquad \mathcal{L}_i \mathcal{U}_i = A^i,$$

that is, i steps of LR applied to A are equivalent to a similarity given by a factorization of A^i.

2.1. Convergence of LR algorithm in the simplest case. We now restrict ourselves to the case when A is nonsingular and has eigenvalues λ_i, $i = 1, \dots, n$, of distinct modulus, so that it necessarily has linear elementary divisors. We may write

$$A = X \operatorname{diag}(\lambda_1, \dots, \lambda_n) X^{-1} =: X \Lambda Y, \tag{2.1}$$

where the columns of X are the right eigenvectors of A and the rows of Y are the row eigenvectors of A.

We order the eigenvalues λ_i, $i = 1, \dots, n$, of A so that they satisfy

$$|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|.$$

Then, under certain mild restrictions, we have

$$L_i \to I \qquad \text{and} \qquad R_i \to A_i \to \begin{bmatrix} \lambda_1 & * & * & * \\ & \lambda_2 & * & * \\ & & \ddots & \vdots \\ & & & \lambda_n \end{bmatrix} \qquad \text{as} \qquad i \to \infty.$$

THEOREM 2.3. *Let $A = X \operatorname{diag}(\lambda_1, \dots, \lambda_n) X^{-1}$. If*

(i) $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$,

(ii) *the leading principal minors of X and X^{-1} are nonzero, i.e., X and $Y = X^{-1}$ admit triangular factorizations*

$$X = L_X U_X \qquad \text{and} \qquad Y = L_Y U_Y,$$

then, for the sequence $\{A_i\}_{i=1}^{\infty}$ *generated by LR algorithm,*

$$\lim A_i = A_\infty$$

exists and is upper triangular with the i^{th} *diagonal entry of* A_∞ *equal to* λ_i*.*

The proof of this theorem is rather technical (see Wilkinson [10] and [11, pp.487-492]) and it was first given by Rutishauser [9]. It is based on the fact that if X in (2.1) admits the triangular factorization

$$X = L_X U_X$$

then

$$\mathcal{L}_i \to L_X \qquad \text{and} \qquad A_i = \mathcal{L}_{i-1}^{-1} A \mathcal{L}_{i-1} \to L_X^{-1} A L_X = U_X X^{-1} A X U_X^{-1} = U_X \operatorname{diag}(\lambda_1, \ldots, \lambda_n) U_X^{-1},$$

showing that the limiting A_i is upper-triangular with diagonal elements $\lambda_1, \ldots, \lambda_n$.

Another important observation is that for a matrix M if

$$M = I + F$$

and $\|F\|$ is sufficiently small, then the triangular decomposition LU of M exists and

$$L \to I,\ U \to I \qquad \text{as} \qquad \|F\| \to 0.$$

The LR method can break down. Such a failure corresponds to the non-existence of a triangular decomposition of a certain matrix $I + G_i$ related to stage i. But since $G_i \to O$ this cannot happen at a late stage in the process.

If we remove the condition that Y has non-vanishing principal minors, a phenomenon called *disorder of latent roots* occurs - A_{i+1} tends to an upper triangular matrix having as its diagonal the eigenvalues but no longer in monotonic decreasing order. The eigenvalues are therefore disordered in the limiting matrix and this phenomenon happens to be unstable in practice.

The speed of convergence of the LR algorithm is determined essentially by the speed at which the quantities $\dfrac{\lambda_k}{\lambda_j}$, $k > j$, tend to zero. If $\left|\dfrac{\lambda_k}{\lambda_j}\right|$ is close to 1, convergence may be slow. So, the speed of convergence of A_i to upper-triangular form depends, particularly, on the ratios $\left|\dfrac{\lambda_{k+1}}{\lambda_k}\right|$ which are the closest to 1.

If we let $L_i = (l_{kj}^{(i)})$, from the relation $L_i = \mathcal{L}_{i-1}^{-1}\mathcal{L}_i$ it can be proved that

$$l_{kj}^{(i)} = \mathcal{O}\left(\frac{\lambda_k}{\lambda_j}\right)^i \qquad \text{as} \qquad i \to \infty \qquad (k > j).$$

So if we have $\lambda_n \approx 0$ then

$$l_{n,j}^{(i)} = \mathcal{O}\left(\frac{\lambda_n}{\lambda_j}\right)^i, \qquad j < n.$$

converges quickly to 0, that is, the last row of L_i converges quickly to $\boldsymbol{e}_n^T$. Thus $A_{i+1} = R_i L_i$ will have the last line converging quickly to $\lambda_n \boldsymbol{e}_n^T \approx \mathbf{0}$. This explains why it is more efficient to apply LR algorithm to $A - \sigma I$ with σ close to λ_n and how shifts of origin accelerate convergence.

Note that in establishing this convergence result we made the assumption that all the eigenvalues are of different magnitude and in this case A can not be a real matrix with complex conjugate eigenvalues. The crucial observation is that if $|\lambda_{j+1}| = |\lambda_j|$ then $|l_{j+1,j}^{(i)}|$ does not diminish as i increases. So we might expect non-convergence when $\lambda_{j+1} = \lambda_j$.

For the case of non-linear divisors a simple counter-example is given in the last section of [10] that shows immediately that the LR algorithm does not necessarily converge to an upper-triangular matrix. The matrix

$$A = \begin{bmatrix} a & 0 & 0 \\ 1 & a & 0 \\ 0 & 1 & a \end{bmatrix}, \quad a \in \mathbb{R},$$

is LR invariant, that is, $A_i = A$ for all i. But a sensible algorithm should detect whether a matrix is triangular (lower or upper) before invoking any eigenvalue finder.

3. Eigenvector properties of an unreduced tridiagonal. Eigenvector matrices of real unreduced symmetric tridiagonal matrices have several attractive properties and have been studied widely in the literature. The eigenvalues are real and distinct and key properties are

- the first and last entries cannot vanish; there are very elegant formulae for the squares of entries of normalized eigenvectors.
- When the off-diagonal entries are all of the same sign, the eigenvector for the rightmost (largest) eigenvalue has no sign changes, for the second largest eigenvalue, the eigenvector has one sign change, and so on. The eigenvector for the leftmost (smallest) eigenvalue has the maximal number of sign changes, namely $n-1$ for a $n \times n$ matrix. See Gantmacher and Krein [2] and Fiedler [1].

Our interest is in the real unsymmetric case and we expect the matrix spectrum to have a mixture of real and complex eigenvalues. Of the properties above only the first extends to our case. The proof is identical to the symmetric case and will be omitted. A new difficulty in our case is that the eigenvalues need not be simple, so the Jordan form may not be diagonal. In such cases the eigenvector matrix must be filled out with the so-called generalized eigenvectors with the property that, for any such C,

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

We say v is an eigenvector of grade j.

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 and Wilkinson for the general case. To the best of our knowledge these results are new.

We will say that X is *strongly* (or *completely*) *regular* with the meaning that X and all its leading principal submatrices are invertible. We shall use the terms “strongly 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 invertible.

We proceed from the easier cases to the more difficult in stages.

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 an unreduced real tridiagonal matrix

$$C = \begin{bmatrix} a_1 & c_1 & & & \\ b_1 & a_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & b_{n-2} & a_{n-1} & c_{n-1} \\ & & & b_{n-1} & a_n \end{bmatrix} \in \mathbb{R}^{n\times n} \tag{3.1}$$

with $b_i c_i \neq 0$, $i = 1, \dots, n-1$.

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.

3.1. All eigenvalues distinct. Suppose all eigenvalues of C are distinct and let the spectrum be

$$\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_n\}.$$

The following matrix plays a key role in our results,

$$P = P_C = [p_{i-1}(\lambda_j)]_{i,j=1}^n \,.$$

The notation means that the (i, j) element of P is equal to $p_{i-1}(\lambda_j)$. So the j^{th} column of P is given by the column vector

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

that is,

$$P = \begin{bmatrix} 1 & 1 & \dots & 1 \\ p_1(\lambda_1) & p_1(\lambda_2) & \dots & p_1(\lambda_n) \\ \vdots & \vdots & & \vdots \\ p_{n-1}(\lambda_1) & p_{n-1}(\lambda_2) & \dots & p_{n-1}(\lambda_n) \end{bmatrix}. \tag{3.2}$$

P and P^T are called *polynomial Vandermonde* matrices. The standard Vandermonde matrix V is defined by

$$V = V_\Lambda = \left[\lambda_i^{j-1}\right]_{i,j=1}^n,$$

that is,

$$V = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \dots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \dots & \lambda_2^{n-1} \\ \vdots & \vdots & & \vdots & \\ 1 & \lambda_n & \lambda_n^2 & \dots & \lambda_n^{n-1} \end{bmatrix}. \tag{3.3}$$

The valuable property of V is that

$$\det(V) = \prod_{i>j} (\lambda_i - \lambda_j) \tag{3.4}$$

where the product extends over all pairs (i, j) with $n \geq i > j \geq 1$; $n(n-1)/2$ terms in all.

When the λ_i are distinct then V is strongly regular because each leading principal submatrix of V is also a Vandermonde built from a subset of the eigenvalues.

Lemma 3.1. *If all the eigenvalues λ_i, $i = 1, \ldots, n$, are distinct then V is strongly regular.*

Moreover ΠV is strongly regular for any permutation matrix Π because no particular ordering of the eigenvalues was specified in the definition.

Lemma 3.2. $\det(P) = \det\left(V^T\right) = \det(V)$.

With the aid of $P = \begin{bmatrix} \boldsymbol{p}(\lambda_1) & \boldsymbol{p}(\lambda_2) & \cdots & \boldsymbol{p}(\lambda_n) \end{bmatrix}$ we can find simple forms for the column and row eigenvectors of C.

Definition 3.3. *Consider matrix C as defined in (3.1). We will denote*

$$D_b := diag(1, b_1, b_1 b_2, b_1 b_2 b_3, \ldots, b_1 b_2 \cdots b_{n-1}), \tag{3.5}$$
$$D_c := diag(1, c_1, c_1 c_2, c_1 c_2 c_3, \ldots, c_1 c_2 \cdots c_{n-1}). \tag{3.6}$$

Theorem 3.4. *With the notation given above*

$$C\left(D_c^{-1} P\right) = \left(D_c^{-1} P\right)\Lambda, \qquad \left(P^T D_b^{-1}\right) C = \Lambda\left(P^T D_b^{-1}\right).$$

Proof. The result is a reformulation of the celebrated three term recurrence (3TR) associated with C,

$$p_{j+1}(\tau) = (\tau - a_{j+1}) p_j(\tau) - b_j c_j p_{j-1}(\tau), \qquad j = 1, 2, \ldots, n-1$$
$$p_1(\tau) = (\tau - a_1) = (\tau - a_1) p_0(\tau), \text{ since } \qquad p_0(\tau) = 1.$$

□

These row and column eigenvectors are not scaled properly to be inverses of each other. Since the row eigenvector for λ_j annihilates all the column eigenvectors for different eigenvalues we may define a special diagonal matrix $\Delta = \Delta_C$ by

Definition 3.5.

$$\left(P^T D_b^{-1}\right)\left(D_c^{-1} P\right) = \Delta := diag(\delta_1, \ldots, \delta_n), \tag{3.7}$$
$$\delta_j = \boldsymbol{p}(\lambda_j)^T \left(D_b D_c\right)^{-1} \boldsymbol{p}(\lambda_j) \neq 0. \tag{3.8}$$

The matrix Δ^{-1} may be attached to either $D_c^{-1} P$ or $D_b^{-1} P$ or shared between them. In general, Δ will be indefinite.

Theorem 3.6. *The spectral decomposition of C with simple eigenvalues may be written*

$$C = \left(D_c^{-1} P\right)\Lambda\left(\Delta^{-1} P^T D_b^{-1}\right) = \left(D_c^{-1} P \Delta^{-1}\right)\Lambda\left(P^T D_b^{-1}\right).$$

Theorem 3.7. *When C has simple eigenvalues then both column and row eigenvector matrices $D_c^{-1} P \Delta^{-1}$ and $P^T D_b^{-1}$, respectively, are strongly regular.*

Next we consider the opposite extreme, a maximal Jordan block of C.

3.2. The one-point spectrum. Suppose now that C's spectrum consists of a single nonzero point λ and such that its Jordan form is

$$J = \lambda I + N$$

where N is the nilpotent matrix

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

We know that

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

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

The most elegant way to find eigenvectors of higher grade is to differentiate the (3TR) as many times as is necessary. Differentiate once,

$$p'_{j+1}(\tau) = (\tau - a_{j+1})p'_j(\tau) + p_j(\tau) - b_j c_j p'_{j-1}(\tau), \qquad j = 1, 2, \ldots, n-1,$$

rearrange the terms and divide by $c_1 \cdots c_j$ to find

$$b_j \frac{p'_{j-1}(\tau)}{c_0 c_1 \cdots c_{j-1}} + (a_{j+1} - \tau)\frac{p'_j(\tau)}{c_0 c_1 \cdots c_j} + c_{j+1}\frac{p'_{j+1}(\tau)}{c_1 \cdots c_{j+1}} = \frac{p_j(\tau)}{c_0 c_1 \cdots c_j}, \qquad j = 0, 1, \ldots, n-1,$$

where we define $p'_{-1}(\tau) = 0$, $b_n = c_n = 1$ and $c_0 = 1$. In matrix terms

$$(C - \tau I)D_c^{-1}\boldsymbol{p}'(\tau) = D_c^{-1}\big(\boldsymbol{p}(\tau) - \boldsymbol{e}_n p'_n(\tau)\big).$$

Hence, when $\tau = \lambda$,

$$(C - \lambda I)D_c^{-1}\boldsymbol{p}'(\lambda) = D_c^{-1}\boldsymbol{p}(\lambda), \qquad \text{since } p'_n(\lambda) = 0,$$

and

$$(C - \lambda I)^2 D_c^{-1}\boldsymbol{p}'(\lambda) = (C - \lambda I)D_c^{-1}\boldsymbol{p}(\lambda) = \boldsymbol{0},$$

since $p_n(\lambda) = (\tau - \lambda)^n$ and λ is a multiple zero of p_n.

To obtain the next vector differentiate the (3TR) twice and divide by 2 to obtain

$$(C - \tau I)\frac{1}{2}D_c^{-1}\boldsymbol{p}''(\tau) = D_c^{-1}\left(\boldsymbol{p}'(\tau) - \boldsymbol{e}_n \frac{1}{2}p''_n(\tau)\right).$$

Putting $\tau = \lambda$ and since $p''(\lambda) = 0$,

$$(C - \lambda I)\frac{1}{2}D_c^{-1}\boldsymbol{p}''(\lambda) = D_c^{-1}\boldsymbol{p}'(\lambda),$$

and

$$(C-\lambda I)^3 \frac{1}{2} D_c^{-1} \boldsymbol{p}''(\lambda) = (C-\lambda I)^2 D_c^{-1} \boldsymbol{p}'(\lambda) = \mathbf{0}.$$

It may be verified that the appropriate definition of P in the confluent case is the unit lower triangular matrix

Definition 3.8.

$$P_\lambda = P = \begin{bmatrix} \boldsymbol{p}(\lambda) & \boldsymbol{p}'(\lambda) & \frac{1}{2!}\boldsymbol{p}''(\lambda) & \ldots & \frac{1}{(n-1)!}\boldsymbol{p}^{(n-1)}(\lambda) \end{bmatrix} \tag{3.9}$$

and so

$$(C-\lambda I) D_c^{-1} P = D_c^{-1} P N. \tag{3.10}$$

A 6×6 example of P is given in section 3.2.1.

So, the matrix of generalized right eigenvectors of C is $D_c^{-1}P$. To find the row eigenvectors for C start from

$$P^T D_b^{-1} C = \left(\lambda I + N^T\right) P^T D_b^{-1} = \not{I}(\lambda I + N)\not{I} P^T D_b^{-1}, \tag{3.11}$$

where, using the notation introduced by G.W. Stewart,

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

is the *reversal* or *anti-diagonal* or *flip* matrix (all entries (i,j) are zero except when $i+j = n+1$). We used the fact that $\not{I} \cdot \not{I} = I$ and $N^T = \not{I} N \not{I}$. Thus, pre-multiplying (3.11) by $\not{I}$, we find that

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

So, $\not{I} P^T D_b^{-1}$ is the matrix of generalized 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, as was the case for simple eigenvalues. The reason is subtle: for a Jordan block, the eigenvectors of grade higher than 1 are not uniquely defined. The phrase "the Jordan basis" that can be found in some text books is incorrect; it is far from unique as we now illustrate.

Consider the equation above in (3.10),

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

Post-multiply by any invertible matrix $\varphi(N)$, φ a polynomial, that comutes with $\lambda I + N$ to find that

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

Thus $D_c^{-1}P$ is only unique up to post-multiplication by any invertible polynomial $\varphi(N)$. And there is no loss in normalizing φ to satisfy $\varphi(O) = I$. However, the preferred choice of $\varphi(N)$ for us is given by

$$\left(\text{Ɨ} P^T D_b^{-1}\right)\left(D_c^{-1}P\right) = \varphi(N). \tag{3.12}$$

We have proved

THEOREM 3.9. *If unreduced tridiagonal C has one-point spectrum λ and P, D_b, D_c are as defined in (3.9), (3.5) and (3.6) then*

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

for a certain polynomial φ with $\varphi(O) = I$, given by (3.12).

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

3.2.1. 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.

Z. S. Liu [4] 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 $\boldsymbol{\gamma}^6 = \begin{bmatrix} -1 & 1 & -1 & 1 & -1 \end{bmatrix}$.

Consider

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

We have

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

Then

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

Now, recall Δ from the case of distinct eigenvalues in definition 3.5,

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

defining $\mathcal{U}$ indirectly. Thus,

$$\mathcal{U} = \left(\text{Ɨ} P^T D_b^{-1} D_c^{-1} P\right)^{-1} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

that is, $\mathcal{U} = I - N^3 + N^5$ is a polynomial in N and thus commutes with $\lambda I + N$.

Thus $\text{Ɨ} P^T D_b^{-1} D_c^{-1} P = \mathcal{U}^{-1} = I + N^3 - N^5 = \varphi(N)$ is unit upper triangular and Toeplitz. Finally, the spectral decomposition of Liu_6^T:

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

Note that $P\mathcal{U}$ is in LU form and $P^T D_b^{-1}$ is upper triangular. It will be shown that $C + \sigma I$, $\sigma \neq 0$, can admit the basic LR algorithm with no breakdowns and will converge to

$$\mathcal{U}(\sigma I + N)\mathcal{U}^{-1} = \sigma I + N.$$

3.3. The general case. In general, unreduced C will have some simple eigenvalues and some multiple ones. The unreduced property implies that C is nonderogatory, meaning that there is only one Jordan block, and thus one eigenvector, for each eigenvalue.

It follows directly from the previous sections that the matrix P has a column $\boldsymbol{p}$ for each simple eigenvalue λ and a block of columns $\left[\boldsymbol{p}(\lambda) \quad \boldsymbol{p}'(\lambda) \quad \frac{1}{2!}\boldsymbol{p}''(\lambda) \quad \ldots \quad \frac{1}{m!}\boldsymbol{p}^{(m)}(\lambda)\right]$ if λ has multiplicity $m+1$. The only constraint on ordering of columns is that each block must be treated as a whole.

As shown in the section on one-point spectrum matrices, the order of the row eigenvectors must be reversed within each block so Ɨ plays a key role.

We state without proof.

Theorem 3.10. *Let J be the (upper) Jordan form of an unreduced matrix C. For D_b and D_c as defined in (3.5) and (3.6), the spectral decomposition may be written*

$$C = D_c^{-1} P\mathcal{U} J \mathcal{W} P^T D_b^{-1}$$

where $\mathcal{U}$ is a unique unit upper triangular matrix that commutes with J and $\mathcal{W}$ is a symmetric permutation matrix that is a direct sum of reversal matrices I conforming to the block structure of J. $\mathcal{W}$ also commutes with J. P has a column for each simple eigenvalue (see (3.2)) and a block of columns for multiple eigenvalues (see (3.9)). Recall P, and therefore P^T, are strongly regular.

4. Convergence of basic LR algorithm on an unreduced tridiagonal. We will show that the assumptions required by Wilkinson in the general case to ensure the convergence of LR algorithm are no longer needed on an unreduced tridiagonal matrix.

The L factor of a matrix M will be denoted by $\mathcal{L}(M)$ and the U factor by $\mathcal{U}(M)$, provided that M is strongly regular. In this new notation we will write

$$C^k = \mathcal{L}(C^k)\mathcal{U}(C^k) \qquad \text{and} \qquad C_{k+1} = \mathcal{L}(C^k)^{-1}C\mathcal{L}(C^k) \tag{4.1}$$

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

4.1. Eigenvalues of distinct absolute value. The result in this section is not entirely new but helps us to understand the new case.

Let C be a nonsingular unreduced tridiagonal matrix. Without loss of generality, we may write

$$\begin{aligned} &C = X\Lambda X^{-1} \\ &\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n), \qquad |\lambda_i| > |\lambda_{i+1}|, \qquad \text{for all } i. \end{aligned}$$

With the notation of the previous section,

$$\begin{aligned} &X = D_c^{-1}P\Delta^{-1}, \qquad X^{-1} = P^T D_b^{-1} \\ &\Delta = P^T D_b^{-1} D_c^{-1} P \quad (= \text{diagonal}) \end{aligned}$$

where $P = P_\Lambda$ is the polynomial Vandermonde matrix given in (3.2). Since P is strongly regular (and also X and X^{-1} by theorem 3.7) we may write

$$P = LDU,$$

where L is unit lower triangular, D is diagonal and U is unit upper triangular.

Next comes the analysis that must be imitated when we turn to the one-point spectrum. We can manipulate C^k into LU form as follows

$$\begin{aligned} C^k &= X\Lambda^k X^{-1} \\ &= D_c^{-1}P\Delta^{-1}\Lambda^k P^T D_b^{-1} \\ &= D_c^{-1}LDU\Delta^{-1}\Lambda^k U^T DL^T D_b^{-1} \\ &= D_c^{-1}LDU\Delta^{-1}(\Lambda^k U^T \Lambda^{-k})\Lambda^k DL^T D_b^{-1} \\ &= (D_c^{-1}LD_c)(D_c^{-1}DU\Delta^{-1})(I + E_k)(\Lambda^k DL^T D_b^{-1}) \\ &= (D_c^{-1}LD_c)(I + F_k)(D_c^{-1}DU\Delta^{-1})(\Lambda^k DL^T D_b^{-1}) \end{aligned} \tag{4.2}$$

where

$$F_k = \left(D_c^{-1}DU\Delta^{-1}\right) E_k \left(D_c^{-1}DU\Delta^{-1}\right)^{-1}.$$

Notice that

$$\begin{aligned} X &= \left(D_c^{-1}LD_c\right)\left(D_c^{-1}DU\Delta^{-1}\right) =: L_X U_X \\ Y &= U^T \left(DL^T D_b^{-1}\right) =: L_Y U_Y \end{aligned} \tag{4.3}$$

are the LU factorizations of X and Y, respectively.

Also, we wrote

$$\Lambda^k U^T \Lambda^{-k} = I + E_k.$$

So, E_k is strictly lower triangular and, if we let $U = (u_{ij})$ and $E_k = \left(e_{ij}^{(k)}\right)$, we have

$$e_{ij}^{(k)} = \begin{cases} u_{ji}\left(\frac{\lambda_i}{\lambda_j}\right)^k, & i > j \\ 0, & i \leq j \end{cases}. \tag{4.4}$$

Thus the $(j+m, j)$ entry of E_k is $u_{j,j+m}\left(\frac{\lambda_{j+m}}{\lambda_j}\right)^k$ and tends to zero as k tends to infinity, since $|\lambda_{j+m}| < |\lambda_j|$. Hence

$$\|F_k\| \leq \operatorname{cond}\left(D_c^{-1} D U \Delta^{-1}\right) \|E_k\| \to 0 \qquad \text{as} \qquad k \to \infty.$$

It seems likely that $I + F_k$ is strongly regular for all k but it certainly holds for large enough k, say

$$I + F_k = L_k U_k \qquad \text{with} \qquad L_k, U_k \to I \qquad \text{as} \qquad k \to \infty.$$

Thus, for large enough k, from (4.2),

$$C^k = \left(D_c^{-1} L D_c\right)\left(L_k U_k\right)\left(D_c^{-1} D U \Delta^{-1}\right)\left(\Lambda^k D L^T D_b^{-1}\right),$$

and then

$$\mathcal{L}(C^k) = D_c^{-1} L D_c L_k$$

since $U_k\left(D_c^{-1} D U \Delta^{-1}\right)\left(\Lambda^k D L^T D_b^{-1}\right)$ is upper triangular. Thus

$$\mathcal{L}(C^k) \to D_c^{-1} L D_c = L_X \qquad \text{as} \qquad k \to \infty.$$

Finally, from (4.1),

$$C_{k+1} = \mathcal{L}(C^k)^{-1} C \mathcal{L}(C^k),$$

and thus

$$\begin{aligned} C_{k+1} &\to \left(D_c^{-1} L D_c\right)^{-1} X \Lambda X^{-1} \left(D_c^{-1} L D_c\right) \\ &= L_X^{-1} L_X U_X \Lambda U_X^{-1} L_X^{-1} L_X \\ &= U_X \Lambda U_X^{-1} \qquad \text{as} \qquad k \to \infty, \end{aligned}$$

showing that C_{k+1} converges to an upper triangular matrix (actually bidiagonal) with diagonal elements equal to $\lambda_1, \ldots, \lambda_n$.

Recalling that the tridiagonal form is preserved by the LR algorithm, we have proved

THEOREM 4.1. *Let C be a nonsingular unreduced tridiagonal matrix with eigenvalues λ_i, $i = 1, \ldots, n$, satisfying $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$. If the LR algorithm does not breakdown, then*

$$\lim C_k = C_\infty$$

exists and is upper bidiagonal with the i^{th} diagonal entry of C_∞ equal to λ_i.

4.2. One-point spectrum. How can the analysis of the distinct absolute value case of the previous section be rebuilt when an eigenvalue is multiple so that no shift will produce different moduli? We will deal first with the case $\lambda \neq 0$.

4.2.1. The case $\boldsymbol{\lambda \neq 0}$. Recall that the Vandermonde matrix P for this case is unit lower triangular,

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

From theorem 3.9,

$$C = X(\lambda I + N)X^{-1} \tag{4.5}$$

with

$$X = D_c^{-1} P \varphi(N)^{-1} \qquad \text{and} \qquad X^{-1} = \not{I} P^T D_b^{-1}$$

where $\varphi(N)$ is given in (3.12). Then

$$C^k = D_c^{-1} P \varphi(N)^{-1} (\lambda I + N)^k \not{I} P^T D_b^{-1}. \tag{4.6}$$

Note that P^T is unit upper triangular.

It is the matrix $\not{I}$ that rescues the algorithm from failure. The following lemma is the key.

LEMMA 4.2. *For all $k \geq n$, $(\lambda I + N)^k \not{I}$ for $\lambda \neq 0$ is strongly regular and thus admits triangular factorization, say*

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

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

This lemma follows from Theorem 4.4 which will be presented later in this section.

THEOREM 4.3. *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 (3.6).

Proof. From (4.6) and with Lemma 4.2 the proof of convergence now imitates the easy case. For $k \geq n$,

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

with

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

since

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

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 $D_c^{-1}\varphi(N)^{-1}U_k P^T D_b^{-1}$ is upper triangular ($\varphi(N)^{-1}$ is upper triangular and Toeplitz).

Finally, since P is unit lower triangular, the LU factorization of X is

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

and then

$$\begin{aligned}
C_{k+1} &= \mathcal{L}(C^k)^{-1} C \mathcal{L}(C^k) \\
&= \mathcal{L}(C^k)^{-1} X(\lambda I + N)X^{-1}\mathcal{L}(C^k) && \text{(by 4.5)} \\
&\quad \to \left(D_c^{-1} P D_c\right)^{-1} X(\lambda I + N)X^{-1}\left(D_c^{-1} P D_c\right) \\
&= L_X^{-1} L_X U_X(\lambda I + N)U_X^{-1} L_X^{-1} L_X \\
&= U_X(\lambda I + N)U_X^{-1} \\
&= D_c^{-1}\varphi(N)^{-1}(\lambda I + N)\varphi(N) D_c \\
&= D_c^{-1}(\lambda I + N) D_c,
\end{aligned}$$

since $\varphi(N)$ commutes with $(\lambda I + N)$. Notice that $D_c^{-1}(\lambda I + N)D_c$ is upper bidiagonal with diagonal entries equal to λ. It is satisfying that $\varphi(N)$ cancels and does not influence the limit. □

Comments on the proof of lemma 4.2. Assume $\lambda \neq 0$. The proof rests on the form of the powers of a Jordan block J. These are Toeplitz and upper triangular matrices and involve binomial coefficients (see [3, p.138]).

For $k \geq n$, since $N^i = O$ for $i \geq n$,

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

where Δ_λ is the matrix defined as

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

So $(\lambda I + N)^k \not{I}$ is a Hankel and upper anti-triangular matrix. Thus, let

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

and then

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

Notice that with a diagonal scaling we are free of the powers of λ. We will now concentrate on H_k and show that H_k has a triangular factorization,

$$H_k = \widetilde{L}_k \widetilde{U}_k.$$

Thus, for each $k \geq n$, $(\lambda I + N)^k \,I\!\!\!/$ also has an LU factorization. We have

$$(\lambda I + N)^k \,I\!\!\!/ = \lambda^k \Delta_\lambda \widetilde{L}_k \widetilde{U}_k \,I\!\!\!/\, \Delta_\lambda^{-1} \,I\!\!\!/ = \Delta_\lambda \widetilde{L}_k \Delta_\lambda^{-1} \lambda^k \Delta_\lambda \widetilde{U}_k \,I\!\!\!/\, \Delta_\lambda^{-1} \,I\!\!\!/ = L_k U_k$$

with $L_k = \Delta_\lambda \widetilde{L}_k \Delta_\lambda^{-1}$ and $U_k = \lambda^k \Delta_\lambda \widetilde{U}_k \,I\!\!\!/\, \Delta_\lambda^{-1} \,I\!\!\!/$. Notice that λ^k is just a scalar converging either to 0 or to ∞ attached to U_k but it does not alter L_k (if $A = LU$ then $\alpha A = L(\alpha U)$). Since $(\lambda I + N)^k \,I\!\!\!/$ is symmetric we can write $L_k U_k = L_k D_k L_k^T$, proving Lemma 4.2. The diagonal matrix D_k is also a function of k but it does not converge to a finite matrix. □

Next we present a result that appears to be new.

For $1 \leq p \leq n$ let $H_n(k)_p$ designate the leading principal $p \times p$ submatrix of H_k. The formulae that follow are the outcome of a difficult determinantal evaluation. We have not found them in the literature but it seems the sort of calculation Sir Thomas Muir would have enjoyed in the 1890's.

Theorem 4.4. *Let $n \in \mathbb{N}$, $1 \leq p \leq n$ 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

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

where

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

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 4.4. 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 4.5. *For $1 \le p \le n$ and $k \ge n$, $H_n(k)_p$ is strongly regular.*

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

Corollary 4.6. *H_k admits triangular factorization*

$$H_k = \widetilde{L}_k \widetilde{U}_k.$$

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 reveals $l_{j+m,j}$ 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 $\mathcal{O}\big((\frac{1}{k})^m\big)$ and thus

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

Also, for $E_k = \Delta_\lambda G_k \Delta_\lambda^{-1}$,

$$L_k \to I + E_k, \;\; E_k \to O \qquad \text{as} \qquad k \to \infty.$$

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

4.2.2. 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.

In exact arithmetic it takes only $2n$ multiplications to determine whether or not C is singular. Suppose that it is singular. It may not be known whether C has a one-point spectrum but well known theory says that if C is singular and permits LU factorization then, in one step, the new transform UL deflates to a matrix of order one less. Thus, if each deflated matrix permits LU factorization, then one would discover the precise multiplicity of 0 as an eigenvalue by forming the LU transform and deflating until there is a nonsingular U.

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 verified 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 left-hand side of the last equation

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

will vanish 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

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

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

[2] 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.

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

[4] 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.

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

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

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

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

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

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

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

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