Chasing the Bulge

Sebastian Gant

5/19/2017

Contents

1 Precursers and Motivation **1**

2 The Reduction to Hessenberg Form **3**

3 The Algorithm **5**

4 Concluding Remarks **8**

5 References **10**

Introduction

In the early days of the computer, numerical analysts faced many difficult tasks that nowadays we take for granted. One such task was that of computing the eigenvalues and eigenvectors of a square matrix over the real or complex numbers. Slow processing speeds and small memories made finding an economic and reliable method for computing eigenvalues a difficult task. In 1959, John Francis formulated and tested his implicitly shifted QR algorithm, which turned the eigenvalue problem into an everyday computation. Versions of his algorithm are used in mathematical computation programs today. This paper is a survey of David S. Watkin's article "Francis's Algorithm" [5], and serves as an introduction to the algorithm.

1 Precursers and Motivation

It is no surprise that the history of eigenvalue computation is closely tied to that of computing the roots of a polynomial. In linear algebra, it is generally taught to factor the characteristic polynomial to find eigenvalues. We have fomulae and simple strategies for finding eigenvalues of $2 \times 2, 3 \times 3$, and 4×4 matrices. But, as Abel proved in 1824, there is no direct method for finding the roots of polynomials of degree 5 or higher. Such methods instead are iterative.

Moreover, zeros of polynomials are very sensitve to small changes in the coefficients, making this method challenging in the field of numerical analysis (this can be seen for zeros of order greater than 1).

Francis knew about the work of Heinz Rutishauser. Under the direction of Eduard Stiefel, Rutishauser set out to solve the following problem: given a matrix A, determine its eigenvalues from a sequence of moments

$$s_k = y^T A^k x$$

for $k = 0, 1, 2, ...$ and arbitrary vectors x, y. The s_ks can be used to define a meromorphic function f by its power series at 0,

$$f(z) = \sum_{k=0}^{\infty} s_k z^k$$

where the poles of f are given exactly by the reciporacles of the eigenvalues of A as follows [1]. Let

$$f(z) = y^T (I - zA)^{-1} x$$

where we define A to be an $n \times n$ symmetric matrix ($A^T = A$) so that A is diagonalizeable and there exists an orthonormal basis of eigenvectors corresponding to real eigenvalues of A. Then $A = Q\Lambda Q^{-1}$ where Λ is the diagonal matrix with eigenvalues λ_i down the diagonal, and Q is the matrix with orthonormal eigenvectors of A in its columns. The matrix Q is thus unitary ($Q^{-1} = Q^T$). If we let $x = y = e_1$, where e_1 is the vector with 1 in its first entry and 0s elsewhere, then

$$f(z) = e_1^T Q(I - z\Lambda)^{-1} Q^T e_1 = \sum_{j=1}^{n} \frac{(q_{1j})^2}{1 - z\lambda_j}$$

where q_{1j} is the first entry of the jth eigenvector of A. Here we can see that the poles of f occur when $z = 1/\lambda_i$.

Rutishauser designed the quotient-difference algorithm to solve this problem (See [2] or [3] for a detailed discussion). But again, the poles of a meromorphic function are sensitive to small changes in the coefficients of f. This eventually led to Rutishauser's formulation of the LR algorithm, on which the QR algorithm is heavily based.

The Basic *QR* Algorithm

John Francis and Vera Kublanovskaya discovered the basic QR algorithm independently. Rutishauser's LR algorithm relies on factoring the matrix into the product of two triangular matrices. The QR algorithm instead factors A into the product of a unitary and an upper trianglular matrix. The following theorem guarentees the existence of such a factorization.

Theorem 1. *Let $A \in M_{n\times n}$. There is unitary matrix Q with orthonormal columns and an upper triangular matrix R such that $A{=}QR$. If A is nonsingular, then R may be chosen so that all of its diagonal entries are positive. In this case, the factorization is unique.*

A proof is given in [4], and the factorization is similar to the Gram-Schmidt process. If A is real, then all of the computations can be carried out in real arithmetic, so that Q and R are real. If we reverse the factors, we have

$$A = QR \qquad A_1 = RQ$$

Where $A_1 = Q^{-1}AQ$. So A is similar to A_1 with Q unitary. Thus, A and A_1 share eigenvalues. We can iterate the process to get

$$A_{k-1} = Q_k R_k \qquad A_k = R_k Q_k$$

Assuming that the eigenvalues are of distinct modulus, the sequence $\{A_k\}$ tends to a limit in (possibly block) upper triangular form, where we can read off the eigenvalues down the diagonal. If the eigenvalues do not have distinct modulus, then cylcing can occur, and the sequence will not tend to a limit. In order to speed up the rate of convergence, we can incorporate shifts in the following way:

$$A_{k-1} - \rho_k I = Q_k R_k \qquad A_k = R_k Q_k + \rho_k I \tag{1}$$

Where I is the identity matrix. Again, the A_ks are all unitarily similar. A good choice in ρ_k is a close approximation to an eigenvalue. So, if we have complex eigenvalues for a real matrix, we would need to perform complex shifts and carry out the operation in complex arithmetic. Alternatively, if we made the first similarity transformation $A \mapsto A_1$ with complex shift ρ, and the second, $A_1 \mapsto A_2$ with $\overline{\rho}$, the resulting matrix A_2 is real again. In the limit, the eigenvalues are revealed in conjugate pairs from the 2×2 blocks down the main diagonal of a block triangular matrix.

In trying to find a way to get directly from A to A_2, Francis discovered his implitly shifted QR algorithm, carried out in real arithmetic. Strangely, the algorithm does not resemble (1) in a direct way. In his classes at Washington State University, Watkins has considered bypassing (1) and teaching Francis's algorithm straightaway [5].

2 The Reduction to Hessenberg Form

Here we introduce some machinary to help us understand and describe Francis's algorithm.

Given two distinct, nonzero vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ with the same Euclidean norm, let $\mathcal{S}$ denote the hyperplane orthogonal to the vector $\mathbf{x} - \mathbf{y}$ that passes through the origin. We can think of a linear transformation Q that maps vectors to the other side of $\mathcal{S}$ like the image of a mirror. Then $\mathbf{x} \mapsto \mathbf{y}$ and $\mathbf{y} \mapsto \mathbf{x}$ under Q. The matrix corresponding to the linear transformation is thus invertible with $Q = Q^{-1}$, and unitary as it preserves norms. We are ready for the following theorem:

Theorem 2. *Let* $\mathbf{x}$ *be a nonzero vector. Then there is a reflector* Q *such that* $Q\mathbf{x} = (\pm\|\mathbf{x}\|_2)\, e_1$.

Surely, given $\mathbf{x}$, let

$$\mathbf{y} = \begin{bmatrix} \pm\|\mathbf{x}\|_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Since $\mathbf{x}$ and $\mathbf{y}$ share the same norm, there is a reflector Q that maps $\mathbf{x}$ to $\mathbf{y}$ and vice versa.

With reflectors, we can perform similarity transformations on a matrix A to put A in what is called **Upper Hessenberg** form. An upper Hessenberg matrix satisfies $a_{ij} = 0$ if $i > j + 1$. This resembles upper triangular form but with one more set of nonzero entries underneath the main diagonal.

Theorem 3. *Every* $A \in M_{n\times n}$ *is orthogonally similar to an upper Hessenberg matrix:*

$$H = Q^{-1}AQ$$

Here we will give an outline of the proof; the details can be found in [6]. We construct Q as the product of $n-2$ reflectors in the following way. Write the first reflector, Q_1, in the form

$$Q_1 = \begin{bmatrix} 1 & \\ & \tilde{Q}_1 \end{bmatrix}$$

where empty entries represent 0s and $\tilde{Q}_1$ is an $(n-1)\times(n-1)$ reflector such that

$$\tilde{Q}_1 \begin{bmatrix} a_{21} \\ a_{31} \\ \vdots \\ a_{n1} \end{bmatrix} = \begin{bmatrix} \alpha \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Then

$$Q_1 A = \begin{bmatrix} 1 & \\ & \tilde{Q}_1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & \\ \vdots & & \ddots & \\ a_{n1} & & & a_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \alpha & * & & \vdots \\ 0 & * & \ddots & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{bmatrix}$$

Where "$*$" denotes a possibly nonzero entry. Note that the matrix on the right has zeros from the third column onwards as desired. All that is left for the

similarity transformation is to multiply on the right by $Q_1^{-1} = Q_1$:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \alpha & * & & \vdots \\ 0 & * & \ddots & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{bmatrix} \begin{bmatrix} 1 & \\ & \tilde{Q}_1 \end{bmatrix} = \begin{bmatrix} a_{11} & * & \cdots & * \\ \alpha & * & & \vdots \\ 0 & * & \ddots & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{bmatrix}$$

which preserves the desired zeros entries. We can repeat this process with another reflector of the form

$$Q_2 = \begin{bmatrix} 1 & & \\ & 1 & \\ & & \tilde{Q}_2 \end{bmatrix}$$

so that $Q_2Q_1AQ_1$ and $Q_2Q_1AQ_1Q_2$ have zeros down the second column from the fourth row onwards. Repeating this process with $n-2$ reflectors, as the last two columns of a Hessenberg Matrix still have nonzero entries in the nth row, we get a matrix of the form

$$Q_{n-2} \cdots Q_2Q_1AQ_1Q_2 \cdots Q_{n-2}$$

Letting $Q = Q_1Q_2 \cdots Q_{n-2}$ we reach the desired result. As Watkins notes, the Hessenberg form is as close as we can get to an upper triangular matrix using a direct method like the one above. This is in accordance with the fact that there is no direct method that gives an explicit solution to the roots of a polynomial of degree 5 or higher. Also note that if we start with an upper Hessenberg matrix in the basic QR algorithm described above, all terms in the sequence would be upper Hessenberg as well, and the matrix multiplication would be much simpler.

3 The Algorithm

Now that we know we can put any matrix into upper Hessenberg form via a set of unitary transformations, suppose A is in such form. Suppose too that $a_{ij} \neq 0$ for $i = j + 1$. If not, we could break up the matrix into two smaller block matrices and begin from there. We willl stick to the case where A is real. To begin, pick m shifts $\rho_1, \rho_2, \ldots, \rho_m$. Again, we want to pick shifts that approximate eigenvalues well. Francis chose $m = 2$ and picked ρ_1, ρ_2 to be the eigenvalues of the 2×2 block matrix in the bottom right corner of A. It could be that this matrix gives complex shifts, in which case we can choose shifts in conjugate pairs, so that the operations are still carried out in real arithmetic. If we consider the polynomial

$$p(z) = (z - \rho_m)(z - \rho_{m-1}) \cdots (z - \rho_1)$$

Then with good choices of ρ_k, $|p(\lambda_k)|$ will be small (where λ_k is the eigenvalue closely approximated by ρ_k. It is hard to make precise the notions of "good/bad"

and "small/large" here, so the reader can consider this to be more of a thought exercise). It follows that m of the $|p(\lambda_k)|$s will be small, while $n-m$ of them will be large. We can index the λ_ks in the following manner:

$$|p(\lambda_n)| \leq |p(\lambda_{n-1})| \leq \cdots \leq |p(\lambda_1)|$$

so that $\lambda_n, \lambda_{n-1}, \ldots, \lambda_{n-m+1}$ are well approximated, while $\lambda_{n-m}, \lambda_{n-m-1}, \ldots, \lambda_n$ are not. It follows that $|p(\lambda_{n-m+1})| < |p(\lambda_{n-m})|$ so that

$$\left| \frac{p(\lambda_{n-m+1})}{p(\lambda_{n-m})} \right| \ll 1$$

Consequently, the entry $a_{n-m+1,n-m}$ in the subdiagonal converges to zero after iterations of Francis's algorithm. The zero in the subdiagonal greatly simplifies the problem, as it allows us to break the matrix into the block form:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

where A_{22} is $m \times m$ and A_{11} is $(n-m) \times (n-m)$ and upper Hessenberg. With $m = 2$ or $m = 3$, the eigenvalues of A_{22} are easy to compute and correspond to eigenvalues of A. We can then repeat the process on the matrix A_{11}. This is the general outline of Francis's algorithm. We will now go into the details.

With shifts $\rho_1, \ldots, \rho_m$ in hand, let

$$p(A) = (A - \rho_m I)(A - \rho_{m-1} I) \cdots (A - \rho_1 I) \tag{2}$$

As we will soon show, we do not actually need the whole matrix $p(A)$, just the first column. Let $\mathbf{x}$ denote this column, then $\mathbf{x} = p(A)e_1$. With A upper Hessenberg, we note that $(A - \rho_1 I)e_1$ has only two non-zero entries in the first two positions, so $(A - \rho_2 I)(A - \rho_1 I)e_1$ has zeros at the 4th entry and following. Thus, $\mathbf{x} = p(A)e_1$ has non zero entries from the $(m+1)$th positions onward. In all cases, $\mathbf{x}$ is real. If we consider the vector $\tilde{\mathbf{x}} \in \mathbb{R}^{m+1}$ consisting of the nonzero entries in $\mathbf{x}$, Theorem 2 guarentees the existence of a reflector $\tilde{Q}_0$ such that $\tilde{Q}_0 \tilde{\mathbf{x}} = \alpha e_1$ where $\alpha = \pm \|\tilde{\mathbf{x}}\|_2$. Now, let Q_0 be the reflector

$$\begin{bmatrix} \tilde{Q}_0 & \\ & I \end{bmatrix}$$

where Q_0 is $n \times n$. Then $Q_0 \mathbf{x} = \alpha^{-1} \mathbf{x}$. Since e_1 picks out the first column of Q_0, this column is proportional to $\mathbf{x}$. To begin a unitary similarity transformation, consider $Q_0^{-1} A = Q_0 A$. Because of the dimension of the block $\tilde{Q}_0$ in Q_0, only the first $m+1$ rows of A are affected. Then, multiplying on the right by Q_0 affects only the first $m+1$ columns. This creates a "bulge" in the original Hessenberg matrix

This can be illustrated with an example. Suppose $n = 6$ and $m = 2$. Then

$$Q_0 A = \begin{bmatrix} * & * & * & & & \\ * & * & * & & & \\ * & * & * & & & \\ & & & 1 & & \\ & & & & 1 & \\ & & & & & 1 \end{bmatrix} \begin{bmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ & * & * & * & * & * \\ & & * & * & * & * \\ & & & * & * & * \\ & & & & * & * \end{bmatrix} = \begin{bmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ & & * & * & * & * \\ & & & * & * & * \\ & & & & * & * \end{bmatrix}$$

And

$$Q_0^{-1} A Q_0 = (Q_0 A) Q_0 = \begin{bmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ & & * & * & * & * \\ & & & * & * & * \\ & & & & * & * \end{bmatrix} \begin{bmatrix} * & * & * & & & \\ * & * & * & & & \\ * & * & * & & & \\ & & & & & \\ & & & & & \\ & & & & & \end{bmatrix} = \begin{bmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ & & & * & * & * \\ & & & & * & * \end{bmatrix}$$

Here we color the bulge red. Note that row $(m + 2)$ now has nonzero entries after this first similarity transformation, while there is a large block of 0s underneath row $(m + 2)$. This is in accordance with the fact that $a_{m+2,m+1} \neq 0$ in the original Hessenberg matrix. Now, by Theorem 3, we are guarenteed a set of unitary similarity transformations to return the matrix $Q_0 A Q_0$ into upper Hessenberg form. For our first bulge chasing matrix,

$$Q_1 = \begin{bmatrix} 1 & \\ & \tilde{Q}_1 \end{bmatrix}$$

For any matrix, we would need $\tilde{Q}_1$ to act on rows 2 through n, but because A has 0s in the fist column after row $(m + 2)$ we only need $\tilde{Q}_1$ to act on rows 2 through $(m + 2)$. Thus, $\tilde{Q}_1$ is an $(m + 1) \times (m + 1)$ reflector such that

$$\tilde{Q}_1 \begin{bmatrix} a_{21} \\ a_{31} \\ \vdots \\ a_{m+2,1} \end{bmatrix} = \begin{bmatrix} \alpha \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Q_1 then has a 1 in the first row and column, an $(m + 1) \times (m + 1)$ matrix in a block on the diagonal $(\tilde{Q}_1)$, and the identity in the bottom right corner:

$$Q_1 = \begin{bmatrix} 1 & & \\ & \tilde{Q}_1 & \\ & & I \end{bmatrix}$$

The product $Q_1 Q_0 A Q_0$ establishes 0s in the first column from row 3 onward. If we multiply again on the right by $Q_1^{-1} = Q_1$ to complete the transformation, the *columns* 2 through $m+2$ are affected, preserving the 0s in the first column as desired. Consequently, the similarity transformation translates the bulge down 1 entry, and 1 entry to the right.

Performing $n - 2$ of these transformations "chases the bulge" down, just below the subdiagonal, and out of the matrix through the bottom right corner. This completes the first iteration of Francis's algorithm. We can see this bulge

chasing with our handy $n = 6, m = 2$ example. In this case, $\tilde{Q}_1$ is a 3×3 matrix, so that

$$Q_1 = \begin{bmatrix} 1 & & & & & \\ & * & * & * & & \\ & * & * & * & & \\ & * & * & * & & \\ & & & & I & \\ & & & & & \end{bmatrix}$$

Then

$$Q_1(Q_0 A Q_0) = \begin{bmatrix} 1 & & & & & \\ & * & * & * & & \\ & * & * & * & & \\ & * & * & * & & \\ & & & & I & \\ & & & & & \end{bmatrix} \begin{bmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ & & & * & * & * \\ & & & & * & * \end{bmatrix} = \begin{bmatrix} * & * & * & * & * & * \\ \alpha & * & * & * & * & * \\ & * & * & * & * & * \\ & * & * & * & * & * \\ & & & * & * & * \\ & & & & * & * \end{bmatrix}$$

And completing the similarity transformation,

$$(Q_1 Q_0 A Q_0) Q_1 = \begin{bmatrix} * & * & * & * & * & * \\ \alpha & * & * & * & * & * \\ & * & * & * & * & * \\ & * & * & * & * & * \\ & & & * & * & * \\ & & & & * & * \end{bmatrix} \begin{bmatrix} 1 & & & & & \\ & * & * & * & & \\ & * & * & * & & \\ & * & * & * & & \\ & & & & I & \\ & & & & & \end{bmatrix} = \begin{bmatrix} * & * & * & * & * & * \\ \alpha & * & * & * & * & * \\ & * & * & * & * & * \\ & * & * & * & * & * \\ & & & & * & * \\ & & & & * & * \end{bmatrix}$$

In this case, 3 more similarity transformation would push the bulge out, resulting in the upper Hessenberg matrix

$$\boldsymbol{A} = Q_4 Q_3 Q_2 Q_1 Q_0 A Q_0 Q_1 Q_2 Q_3 Q_4 = \boldsymbol{Q}^{-1} A \boldsymbol{Q}$$

for $\boldsymbol{Q} = Q_0 Q_1 Q_2 Q_3 Q_4$. Watkins [5] advocates for renaming what is commonly known as the "Implicit QR algorithm" to "Francis's algorithm" because there are no QR decompositions. The common name is misleading in this way. For the next iteration, we can pick some possibly different shifts $\rho_1, \ldots, \rho_m$. Again, with good choices in our ρ_ks, we can get more rapid convergence of $a_{n-m+1,n-m} \to 0$. This allows us to "deflate" the problem into solving for the eigenvalues of two block matrices separately, one of which is $m \times m$, the other $(n-m) \times (n-m)$. With small $m = 2$ or $m = 3$, we get two or three eigenvalues easily.

4 Concluding Remarks

Subspace iteration lies at the core of Francis's algorithm [5]. If we pick a k-dimensional subspace $\mathcal{S}$ of $\mathbb{R}^n$, and build a sequence of subspaces through multiplication by the matrix A

$$\mathcal{S}, A\mathcal{S}, A^2\mathcal{S}, \ldots$$

Where $A^j\mathcal{S} = \{A^j \mathbf{x} : \mathbf{x} \in \mathcal{S}\}$. For simplicity, suppose all of the terms in the sequence have dimension k. We also assume that A is diagonalizeable with n

linearly independent eigenvectors $v_1, v_2, \ldots, v_n$ $(n > k)$. Sort the corresponding eigenvalues so that $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$. So for any vector $\mathbf{x} \in \mathcal{S}$,

$$\mathbf{x} = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$$

where $c_j \in \mathbb{R}$. And because A is linear,

$$A^j \mathbf{x} = c_1 \lambda_1^j v_1 + \cdots + c_k \lambda_k^j v_k + c_{k+1} \lambda_{k+1}^j v_{k+1} + \ldots c_n \lambda_n^j v_n$$

Now, if $|\lambda_k| > |\lambda_{k+1}|$, the components $c_1 \lambda_1^j + \cdots + c_k \lambda_k^j$ in directions $v_1, \ldots, v_k$ will grow much faster than the components in directions $v_{k+1}, \ldots, v_n$. As a result, the sequence above will converge to the k-dimensional subspace spanned by $v_1, \ldots, v_k$. But if $|\lambda_k|$ is not much larger than $|\lambda_{k+1}|$, so that the ratio $|\lambda_{k+1}|/|\lambda_k|$ is close to one, the convergence is somewhat slow. We can replace the matrix A by $p(A)$ as in (2), which is a polynomial of degree m. Now the rate of convergence, is dependent on the ratio $|p(\lambda_{k+1})|/|p(\lambda_k)|$. Again, this ratio is small with shifts $\rho_1, \ldots, \rho_m$ that approximate eigenvalues closely.

Moreover, we note that the "bulge creating" matrix Q_0 was created so that

$$Q_0 e_1 = \frac{1}{\alpha}\mathbf{x} = \frac{1}{\alpha} p(A) e_1$$

Each of the Q_ks for $k = 1, 2, \ldots, n-2$ were constructed to have e_1 in the first column. Denoting $\boldsymbol{Q} = Q_0 Q_1 \cdots Q_{n-2}$ as above, we have

$$\boldsymbol{Q} e_1 = Q_0 Q_1 \cdots Q_{n-2} e_1 = Q_0 e_1 = \frac{1}{\alpha} p(A) e_1$$

so that the first column of $\boldsymbol{Q}$ is proportional to the first column of $p(A)$. This saves a lot of computation, as we only needed to calculate the first column of $p(A)$, which is m matrix-vector products where the matrices $(A - \rho_k I)$ are upper Hessenberg.

As we have seen, the operations carried out through iterations of Francis's algorithm are economical because A is upper Hessenberg. Working with these special matrices has made Francis's algorithm competitive and relevant almost 60 years after its formulation.

5 References

1. G.H. Golub, G. Meurant. (2010). *Matrices, Moments and Quadrature with Applications*, pp. 36-38. Princeton, New Jersey: Princeton University Press.

2. M. Gutknecht and B. Parlett. From qd to LR, or, How were the qd and LR algorithms discovered? *IMA Journal of Numerical Analysis*, 2009. http://www.sam.math.ethz.ch/ mhg/pub/GutPar-IMA.pdf. Retrieved 16 May 2017.

3. P. Henrici. (1974). *Applied and Computational Complex Analysis, Vol. 1*, pp. 608-621. New York, New York: John WIley & Sons, Inc.

4. R. Horn, C. Johnson. (1985). *Matrix Analysis*, pp. 112-115. New York, New York: Cambridge University Press.

5. D. S. Watkins. Francis's Algorithm. *the American Mathematical Monthly*, vol. 118, no. 5, 2011, pp. 387-403. *JSTOR* www.jstor.org/stable/ 10.4169/amer.math.monthly.118.05.387. Retrieved 15 May 2017.

6. D.S. Watkins. (2002). *Fundamental Concepts of Matrix Computations 2nd ed.*, pp. 350-351. New York, New York: John Wiley & Sons, Inc.