

What Is a Companion Matrix?

Nicholas J. Higham*

March 23, 2021

A companion matrix $C \in \mathbb{C}^{n \times n}$ is an upper Hessenberg matrix of the form

$$C = \begin{bmatrix} a_{n-1} & a_{n-2} & \dots & \dots & a_0 \\ 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & & \ddots & 0 & 0 \\ 0 & \dots & \dots & 1 & 0 \end{bmatrix}.$$

Alternatively, C can be transposed and permuted so that the coefficients a_i appear in the first or last column or the last row. By expanding the determinant about the first row it can be seen that

$$\det(\lambda I - C) = \lambda^n - a_{n-1}\lambda^{n-1} - \dots - a_1\lambda - a_0, \quad (*)$$

so the coefficients in the first row of C are the coefficients of its characteristic polynomial. (Alternatively, in $\lambda I - C$ add λ^{n-j} times the j th column to the last column for $j = 1 : n-1$, to obtain $p(\lambda)e_1$ as the new last column, and expand the determinant about the last column.) MacDuffee (1946) introduced the term “companion matrix” as a translation from the German “Begleitmatrix”.

Setting $\lambda = 0$ in $(*)$ gives $\det(C) = (-1)^{n+1}a_0$, so C is nonsingular if and only if $a_0 \neq 0$. The inverse is

$$C^{-1} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ \frac{1}{a_0} & -\frac{a_{n-1}}{a_0} & -\frac{a_{n-2}}{a_0} & \dots & -\frac{a_1}{a_0} \end{bmatrix}.$$

Note that $P^{-1}C^{-1}P$ is in companion form, where $P = I_n(n : -1 : 1, :)$ is the reverse identity matrix, and the coefficients are those of the polynomial $-\lambda^n p(1/\lambda)$, whose roots are the reciprocals of those of p .

*Department of Mathematics, University of Manchester, Manchester, M13 9PL, UK (nick.higham@manchester.ac.uk).

A companion matrix has some low rank structure. It can be expressed as a unitary matrix plus a rank-1 matrix:

$$C = \begin{bmatrix} 0 & 0 & \dots & \dots & 1 \\ 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & \ddots & & 0 \\ \vdots & & \ddots & 0 & 0 \\ 0 & \dots & \dots & 1 & 0 \end{bmatrix} + e_1 [a_{n-1} \ a_{n-2} \ \dots \ a_0 - 1]. \quad (1)$$

Also, C^{-T} differs from C in just the first and last columns, so $C^{-T} = C + E$, where E is a rank-2 matrix.

If λ is an eigenvalue of C then $[\lambda^{n-1}, \lambda^{n-2}, \dots, \lambda, 1]^T$ is a corresponding eigenvector. The last $n - 1$ rows of $\lambda I - C$ are clearly linearly independent for any λ , which implies that C is nonderogatory, that is, no two Jordan blocks in the Jordan canonical form contain the same eigenvalue. In other words, the characteristic polynomial is the same as the minimal polynomial.

The MATLAB function `compan` takes as input a vector $[p_1, p_2, \dots, p_{n+1}]$ of the coefficients of a polynomial, $p_1 x^n + p_2 x^{n-1} + \dots + p_n x + p_{n+1}$, and returns the companion matrix with $a_{n-1} = -p_2/p_1, \dots, a_0 = -p_{n+1}/p_1$.

Perhaps surprisingly, the singular values of C have simple representations, found by Kenney and Laub (1988):

$$\begin{aligned} \sigma_1^2 &= \frac{1}{2} \left(\alpha + \sqrt{\alpha^2 - 4a_0^2} \right), \\ \sigma_i^2 &= 1, \quad i = 2: n-1, \\ \sigma_n^2 &= \frac{1}{2} \left(\alpha - \sqrt{\alpha^2 - 4a_0^2} \right), \end{aligned}$$

where $\alpha = 1 + a_0^2 + \dots + a_{n-1}^2$. These formulae generalize to block companion matrices, as shown by Higham and Tisseur (2003).

Applications

Companion matrices arise naturally when we convert a high order difference equation or differential equation to first order. For example, consider the Fibonacci numbers 1, 1, 2, 3, 5, \dots , which satisfy the recurrence $f_n = f_{n-1} + f_{n-2}$ for $n \geq 2$, with $f_0 = f_1 = 1$. We can write

$$\begin{bmatrix} f_n \\ f_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} f_{n-1} \\ f_{n-2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^2 \begin{bmatrix} f_{n-2} \\ f_{n-3} \end{bmatrix} = \dots = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{n-1} \begin{bmatrix} f_1 \\ f_0 \end{bmatrix},$$

where $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ is a companion matrix. This expression can be used to compute f_n in $O(\log_2 n)$ operations by computing the matrix power using binary powering.

As another example, consider the differential equation

$$y''' = b_2 y'' + b_1 y' + b_0 y.$$

Define new variables

$$z_1 = y'', \quad z_2 = y', \quad z_3 = y.$$

Then

$$\begin{aligned} z'_1 &= b_2 z_1 + b_1 z_2 + b_0 z_3, \\ z'_2 &= z_1 \\ z'_3 &= z_2, \end{aligned}$$

or

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}' = \begin{bmatrix} b_2 & b_1 & b_0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix},$$

so the third order scalar equation has been converted into a first order system with a companion matrix as coefficient matrix.

Computing Polynomial Roots

The MATLAB function `roots` takes as input a vector of the coefficients of a polynomial and returns the roots of the polynomial. It computes the eigenvalues of the companion matrix associated with the polynomial using the `eig` function. As Moler (1991) explained, MATLAB used this approach starting from the first version of MATLAB, but it does not take advantage of the structure of the companion matrix, requiring $O(n^3)$ flops and $O(n^2)$ storage instead of the $O(n^2)$ flops and $O(n)$ storage that should be possible given the structure of C . Since the early 2000s much research has aimed at deriving methods that achieve this objective, but numerically stable methods proved elusive. Finally, a backward stable algorithm requiring $O(n^2)$ flops and $O(n)$ storage was developed by Aurentz, Mach, Vandebril, and Watkins (2015). It uses the QR algorithm and exploits the unitary plus low rank structure shown in (1). Here, backward stability means that the computed roots are the eigenvalues of $C + \Delta C$ for some ΔC with $\|\Delta C\| \leq c_n u \|C\|$. It is not necessarily the case that the computed roots are the exact roots of a polynomial with coefficients $a_i + \Delta a_i$ with $|\Delta a_i| \leq c_n u \max_i |a_i|$ for all i .

Rational Canonical Form

It is an interesting observation that

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -a_3 \\ 0 & 1 & -a_3 & -a_2 \\ 1 & -a_3 & -a_2 & -a_1 \end{bmatrix} \begin{bmatrix} a_3 & a_2 & a_1 & a_0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & -a_3 & 0 \\ 1 & -a_3 & -a_2 & 0 \\ 0 & 0 & 0 & a_0 \end{bmatrix}.$$

Multiplying by the inverse of the matrix on the left we express the 4×4 companion matrix as the product of two symmetric matrices. The obvious generalization of this factorization to $n \times n$ matrices shows that we can write

$$C = S_1 S_2, \quad S_1 = S_1^T, \quad S_2 = S_2^T. \quad (2)$$

We need the rational canonical form of a matrix, described in the next theorem, which Halmos (1991) calls “the deepest theorem of linear algebra”. Let \mathbb{F} denote the field \mathbb{R} or \mathbb{C} .

Theorem 1 (rational canonical form). *If $A \in \mathbb{F}^{n \times n}$ then $A = X^{-1}CX$ where $X \in \mathbb{F}^{n \times n}$ is nonsingular and $C = \text{diag}(C_i) \in \mathbb{F}^{n \times n}$, with each C_i a companion matrix.*

The theorem says that every matrix is similar over the underlying field to a block diagonal matrix composed of companion matrices. Since we do not need it, we have omitted from the statement of the theorem the description of the C_i in terms of the irreducible factors of the characteristic polynomial. Combining the factorization (2) and Theorem 1 we obtain

$$\begin{aligned} A &= X^{-1}CX = X^{-1}S_1S_2X \\ &= X^{-1}S_1X^{-T} \cdot X^TS_2X \\ &\equiv \tilde{S}_1\tilde{S}_2, \quad \tilde{S}_1 = \tilde{S}_1^T, \quad \tilde{S}_2 = \tilde{S}_2^T. \end{aligned}$$

Since S_1 is nonsingular, and since S_2 can alternatively be taken nonsingular by considering the factorization of A^T , this proves a theorem of Frobenius.

Theorem 2 (Frobenius, 1910). *For any $A \in \mathbb{F}^{n \times n}$ there exist symmetric $S_1, S_2 \in \mathbb{F}^{n \times n}$, either one of which can be taken nonsingular, such that $A = S_1S_2$.*

Note that if $A = S_1S_2$ with the S_i symmetric then $AS_1 = S_1S_2S_1 = S_1A^T = (AS_1)^T$, so AS_1 is symmetric. Likewise, S_2A is symmetric.

Factorization

Fiedler (2003) noted that a companion matrix can be factorized into the product of n simpler factors, $n - 1$ of them being the identity matrix with a 2×2 block placed on the diagonal, and he used this factorization to determine a matrix \tilde{C} similar to C . For $n = 5$ it is

$$\tilde{C} = \begin{bmatrix} a_4 & a_3 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & a_2 & 0 & a_1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_0 & 0 \end{bmatrix} = \left[\begin{array}{cc|cc|c} a_4 & 1 & & & \\ 1 & 0 & & & \\ \hline & & a_2 & 1 & \\ & & 1 & 0 & \\ \hline & & & & a_0 \end{array} \right] \left[\begin{array}{c|cc|c} 1 & & & \\ \hline & a_3 & 1 & \\ & 1 & 0 & \\ \hline & & & a_1 & 1 \\ & & & 1 & 0 \end{array} \right].$$

In general, Fielder's construction yields an $n \times n$ pentadiagonal matrix \tilde{C} that is not simply a permutation similarity of C . The fact that \tilde{C} has block diagonal factors opens the possibility of obtaining new methods for finding the eigenvalues of C . This line of research has been extensively pursued in the context of polynomial eigenvalue problems (see Mackey, 2013).

Generalizations

The companion matrix is associated with the monomial basis representation of the characteristic polynomial. Other polynomial bases can be used, notably orthogonal polynomials, and this leads to generalizations of the companion matrix having coefficients on the main diagonal and the subdiagonal and superdiagonal. Good (1961) calls the

matrix resulting from the Chebyshev basis a *colleague matrix*. Barnett (1981) calls the matrices corresponding to orthogonal polynomials *comrade matrices*, and for a general polynomial basis he uses the term *confederate matrices*. Generalizations of the properties of companion matrices can be derived for these classes of matrices.

Bounds for Polynomial Roots

Since the roots of a polynomial are the eigenvalues of the associated companion matrix, or a Fiedler matrix similar to it, or indeed the associated comrade matrix or confederate matrix, one can obtain bounds on the roots by applying any available bounds for matrix eigenvalues. For example, since any eigenvalue λ of matrix A satisfies $|\lambda| \leq \|A\|$, by taking the 1-norm and the ∞ -norm of the companion matrix C we find that any root λ of the polynomial (*) satisfies

$$|\lambda| \leq \max(|a_0|, 1 + \max_{j=1:n-1} |a_j|),$$

$$|\lambda| \leq \max(1, |a_{n-1}| + |a_{n-2}| + \cdots + |a_0|),$$

either of which can be the smaller. A rich variety of such bounds is available, and these techniques extend to matrix polynomials and the corresponding block companion matrices.

References

This is a minimal set of references, which contain further useful references within.

- Jared L. Aurentz, Thomas Mach, Raf Vandebril, and David S. Watkins, Fast and Backward Stable Computation of Roots of Polynomials, SIAM J. Matrix Anal. Appl. 36(3), 942–973, 2015.
- Stephen Barnett, Congenial Matrices, Linear Algebra Appl. 41, 277–298, 1981.
- Fernando De Terán and Froilán M. Dopico and Javier Pérez, New Bounds for Roots of Polynomials Based on Fiedler Companion Matrices, Linear Algebra Appl. 451, 197–230, 2014.
- Miroslav Fiedler, A Note on Companion Matrices, Linear Algebra Appl. 372, 325–331, 2003.
- Nicholas J. Higham and Françoise Tisseur, Bounds for Eigenvalues of Matrix Polynomials, Linear Algebra Appl. 358, 5–22, 2003.
- Charles S. Kenney and Alan J. Laub, Controllability and Stability Radii for Companion Form Systems, Math. Control Signals Systems 1, 239–256, 1988.
- Cyrus Colton MacDuffee, The Theory of Matrices, Chelsea, New York, 1946.
- D. Steven Mackey, The Continuing Influence of Fiedler’s Work on Companion Matrices, Linear Algebra Appl. 439, 810–817, 2013.
- Cleve Moler, ROOTS—of Polynomials, That Is, The MathWorks Newsletter 5(1), 1991.
- Olga Taussky, The Role of Symmetric Matrices in the Study of General Matrices, Linear Algebra Appl. 5, 147–154, 1972.

Related Blog Posts

- What Is a Block Matrix? (2020)
- What Is the Singular Value Decomposition? (2020)

This article is part of the “What Is” series, available from <https://nhigham.com/category/what-is> and in PDF form from the GitHub repository <https://github.com/nhigham/what-is>.