DE GRUYTER OPEN Spec. Matrices 2014; 2:99–105

Special Matrices
Research Article Open Access

Minghua Lin and Harald K. Wimmer*

Bilinear characterizations of companion matrices

Abstract: Companion matrices of the second type are characterized by properties that involve bilinear maps.

Keywords: Companion matrix; reachability matrix; bilinear map

MSC: 15A03, 15A15, 93B05

DOI 10.2478/spma-2014-0010
Received December 14, 2013; accepted May 7, 2014.

1 Introduction

Let K be a field. The matrix

$$F_p = \begin{bmatrix} 0 & \cdot & & & p_0 \\ 1 & 0 & \cdot & & p_1 \\ & \cdot & \cdot & \cdot & \\ & & \cdot & \cdot & \cdot \\ & & & \cdot & \cdot \quad \cdot \\ 0 & 0 & & 1 & p_{n-1} \end{bmatrix} \tag{1.1}$$

is the *second companion matrix* or (in the terminology of [1]) the *companion matrix of the second type* associated with

$$p = [p_0, p_1, \ldots, p_{n-1}]^T \in K^n.$$

It is well known (see e.g. [5] for references) that companion matrices are important in linear algebra, numerical analysis and applications, e.g. in systems and control theory and signal processing. In this paper we focus on bilinear properties of companion matrices that play a role in the single-input case of sensor-only fault detection and identification [4], [2].

Let $A \in K^{n \times n}$, and $b = [b_0, b_1, \ldots, b_{n-1}]^T \in K^n$, $g \in K^n$. Set $e_{n-1} = [0, \ldots, 0, 1]^T \in K^n$ such that $F_p e_{n-1} = p$. The maps

$$h : K^n \times K^n \to K^n, \ (b, g) \mapsto h(A; b, g) = [b, A, \ldots, A^{n-1}b]g,$$

and

$$u : K^n \times K^n \to K^n, \quad (b, g) \mapsto u(A; b, g) = \begin{bmatrix} e_{n-1}^T(b_0 I_n + b_1 A + b_2 A^2 + \cdots b_{n-1} A^{n-1}) \\ e_{n-1}^T(b_1 I_n + b_2 A + \cdots b_{n-1} A^{n-2}) \\ \vdots \\ e_{n-1}^T(b_{n-2} I_n + b_{n-1} A) \\ e_{n-1}^T(b_{n-1} I_n) \end{bmatrix} g \tag{1.2}$$

Minghua Lin: Department of Mathematics and Statistics, University of Victoria, BC, Canada, V8W 3R4, E-mail: mlin87@ymail.com
***Corresponding Author: Harald K. Wimmer:** Mathematisches Institut, Universität Würzburg, 97074 Würzburg, Germany, E-mail: wimmer@mathematik.uni-wuerzburg.de

are bilinear. It is the purpose of this note to show that a matrix A is a second companion matrix if and only if the maps $h(A; b, g)$ and $u(A; b, g)$, respectively, are symmetric. These results will be proved in Section 2. In Section 3 we deal with block companion matrices and we describe extensions of results of Section 2 to matrix polynomials.

2 The main results

We shall use the following notation. The vectors

$$e_0 = [1, 0, \dots, 0]^T, \; \dots, \; e_{n-1} = [0, \dots, 0, 1]^T,$$

are the unit vectors of K^n. The characteristic polynomial of a matrix $A \in K^{n\times n}$ will be denoted by $\chi_A(z)$. If $g \in K^n$ then the Krylov matrix

$$R(A, g) = [g, Ag, \dots, A^{n-1}g] \in K^{n\times n}$$

is the *reachability matrix* (see e.g. [9]) of the pair (A, g). Note that

$$R(A, b)g = h(A; b, g).$$

We recall [9] that A is similar to a companion matrix if and only if $R(A, g)$ has full rank for some $g \in K^n$. With a vector $b = [b_0, b_1, \dots, b_{n-1}]^T \in K^n$ we associate the polynomial $b(z) = b_0 + b_1 z + \cdots + b_{n-1}z^{n-1}$. Thus $b(A) = \sum_{i=0}^{n-1} b_i A^i$. In particular, $e_0(z) = 1, \dots, e_{n-1}(z) = z^{n-1}$. Moreover, if $A \in K^{n\times n}$ then

$$R(A, g)b = b(A)g. \tag{2.1}$$

The following lemma [6] characterizes companion matrices in terms of reachability matrices. To make our note self-contained we include a proof.

Lemma 2.1. *For a matrix $A \in K^{n\times n}$ the following statements are equivalent.*

(i) $A = F_p$ *for some* $p \in K^n$.
(ii) $R(A, g) = g(A)$ *for all* $g \in K^n$.
(iii) $R(A, e_0) = e_0(A) = I_n$.

Proof. It is obvious that A is a companion matrix of the form (1.1) if and only if

$$A[e_0, e_1, \dots, e_{n-2}] = [e_1, e_2, \dots, e_{n-1}]. \tag{2.2}$$

(i) ⇒ (ii): We have to show that $R(F_p, g) = g(F_p)$ holds for all $g = \sum_{i=0}^{n-1} g_i e_i$. From

$$e_i = F_p^i e_0, \;\; i = 0, \dots, n-1, \tag{2.3}$$

follows $R(F_p, e_0) = I_n$. Therefore

$$R(F_p, e_i) = F_p^i\, R(F_p, e_0)$$

implies $R(F_p, e_i) = F_p^i$. Hence

$$R(F_p, g) = \sum_{i=0}^{n-1} g_i\, R(F_p, e_i) = \sum_{i=0}^{n-1} g_i F_p^i = g(F_p).$$

The implication (ii) ⇒ (iii) is obvious.

(iii) ⇒ (i): From $R(A, e_0) = I_n$ follows (2.2). Therefore, A is a companion matrix. □

It was shown in [3] and [4, Proposition A.2] that matrices in second companion form satisfy an identity with “curiously commuting vectors”, stated in (2.4) below. We note that the identity (2.4) is equivalent to the symmetry of the map $h(A; b, g)$.

Theorem 2.2. *Let $A \in K^{n \times n}$. We have $A = F_p$ for some $p \in K^n$ if and only if*

$$[g, Ag, \dots, A^{n-1}g]b = [b, Ab, \dots, A^{n-1}b]g. \tag{2.4}$$

Proof. According to Lemma 2.1 we have $A = F_p$ if and only if $b(A) = R(A, b)$ for all $b \in K^n$, that is,

$$b(A)g = R(A, b)g \text{ for all } b, g \in K^n. \tag{2.5}$$

Then (2.1) implies that (2.5) is equivalent to $R(A, b)g = R(A, g)b$ for all $g, b \in K^n$. □

We now deal with the map $u(A; b, g)$.

Theorem 2.3. *Let $A \in K^{n \times n}$. The following statements are equivalent.*

(i) *$A = F_p$ for some $p \in K^n$.*
(ii) *We have*

$$\begin{bmatrix} e_{n-1}^T A^{n-1} \\ e_{n-1}^T A^{n-2} \\ \vdots \\ . \\ e_{n-1}^T A \\ e_{n-1}^T \end{bmatrix} = \begin{bmatrix} 1 & e_{n-1}^T A e_{n-1} & \dots & & e_{n-1}^T A^{n-1} e_{n-1} \\ & 1 & \dots & & e_{n-1}^T A^{n-2} e_{n-1} \\ & & \ddots & & \vdots \\ & & & 1 & e_{n-1}^T A e_{n-1} \\ & & & & 1 \end{bmatrix}. \tag{2.6}$$

(iii) *The bilinear map $u(A; b, g)$ in (1.2) satisfies $u(A; b, g) = u(A; g, b)$ for all $b, g \in K^n$.*

Proof. Define

$$L(A, g) = \begin{bmatrix} e_{n-1}^T g & e_{n-1}^T Ag & \dots & & e_{n-1}^T A^{n-1} g \\ & e_{n-1}^T g & e_{n-1}^T Ag & \dots & e_{n-1}^T A^{n-2} g \\ & & \ddots & \ddots & \vdots \\ & & & e_{n-1}^T g & e_{n-1}^T Ag \\ & & & & e_{n-1}^T g \end{bmatrix}$$

and

$$Q(A, g) = \begin{bmatrix} e_{n-1}^T (g_0 I_n + g_1 A + g_2 A^2 + \cdots g_{n-1} A^{n-1}) \\ e_{n-1}^T (g_1 I_n + g_2 A + \cdots g_{n-1} A^{n-2}) \\ \vdots \\ e_{n-1}^T (g_{n-2} I_n + g_{n-1} A) \\ e_{n-1}^T (g_{n-1} I_n) \end{bmatrix}.$$

Then $u(A; b, g)$ in (1.2) can be written as

$$u(A; b, g) = L(A, g)b.$$

We have $u(A; g, b) = Q(A, g)b$. Hence (iii) holds if and only if

$$L(A, g) = Q(A, g) \text{ for all } g \in K^n. \tag{2.7}$$

The maps $g \mapsto L(A, g)$ and $g \mapsto Q(A, g)$ are linear. Hence (2.7) is equivalent to

$$L(A, e_i) = Q(A, e_i), \ i = 0, \dots, n-1. \tag{2.8}$$

In the case $i = n - 1$ the corresponding matrices are

$$Q(A, e_{n-1}) = \begin{bmatrix} e_{n-1}^T A^{n-1} \\ e_{n-1}^T A^{n-2} \\ \vdots \\ e_{n-1}^T A \\ e_{n-1}^T \end{bmatrix}$$

and

$$L(A, e_{n-1}) = \begin{bmatrix} 1 & e_{n-1}^T A e_{n-1} & \dots & & e_{n-1}^T A^{n-1} e_{n-1} \\ & 1 & e_{n-1}^T A e_{n-1} & \dots & e_{n-1}^T A^{n-2} e_{n-1} \\ & & \ddots & \ddots & \vdots \\ & & & 1 & e_{n-1}^T A e_{n-1} \\ & & & & 1 \end{bmatrix}.$$

(i) $\Rightarrow$ (iii): Suppose $A = F_p$. Then $A^s e_i = A^{s+i} e_0$ if $0 \le i \le n-1$ and $s \in \mathbb{N}_0$. We have

$$L(A, g)e_r = [e_{n-1}^T A^r g, \dots, e_{n-1}^T A g, e_{n-1}^T g, 0, \dots, 0]^T.$$

Thus

$$L(A, e_i)e_r = [e_{n-1}^T A^{r+i} e_0, \dots, e_{n-1}^T A^{i+1} e_0, e_{n-1}^T A^i e_0, 0, \dots, 0]^T.$$

From

$$Q(A, e_i) = \begin{bmatrix} e_{n-1}^T A^i \\ \vdots \\ e_{n-1}^T A \\ e_{n-1}^T \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

we obtain $Q(A, e_i)e_r = L(A, e_i)e_r$, $r = 0, \dots, n-1$. Hence (2.8) is satisfied, which is equivalent to (iii).

(iii) $\Rightarrow$ (ii): We have seen that (iii) is equivalent to (2.8). Choosing $i = n-1$ in (2.8) yields $Q(A, e_{n-1}) = L(A, e_{n-1})$, that is (2.6).

(ii) $\Rightarrow$ (i): Suppose (2.6) holds. Let

$$A^k = \left[a_{ij}^{(k)}\right]_{i,j=0}^{n-1}, \quad k = 0, 1, \dots, n-1.$$

We show that the rows of A are the rows of a companion matrix. The proof is by induction. The induction hypothesis is

$$e_{n-\nu}^T A = e_{n-\nu-1}^T + a_{n-\nu,n-1} e_{n-1}^T = [0, \dots, 0, 1, 0, \dots, 0, a_{n-\nu,n-1}], \quad 1 \le \nu < n-1. \tag{2.9}$$

From (2.6) we obtain

$$e_{n-1}^T A = [0, \dots, 0, 1, e_{n-1}^T A e_{n-1}] = [0, \dots, 0, 1, a_{n-1,n-1}].$$

Hence (2.9) is satisfied for $\nu = 1$. Assume that (2.9) is valid for $\nu = 1, \dots, k$. Then $A = \begin{bmatrix} W_k \\ H_k \end{bmatrix}$ with

$$\begin{bmatrix} e_{n-k}^T \\ \vdots \\ e_{n-2}^T \\ e_{n-1}^T \end{bmatrix} A = H_k = \begin{bmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 & a_{n-k,n-1} \\ & & & & \ddots & & & \\ 0 & \dots & 0 & 0 & \dots & 1 & 0 & a_{n-2,n-1} \\ 0 & \dots & 0 & 0 & \dots & 0 & 1 & a_{n-1,n-1} \end{bmatrix}_{k \times n}. \tag{2.10}$$

Set

$$e_{n-1-k}^T A = [\tilde{a}_0, \dots, \tilde{a}_{n-2}, \tilde{a}_{n-1}].$$

From (2.6) and (2.10) follows

$$e_{n-1}^T A^{k+1} = (e_{n-1}^T A^k)A = [0, \dots, 0, 1, a_{n-1,n-1}, a_{n-1,n-1}^{(2)}, \dots, a_{n-1,n-1}^{(k)}]$$

$$\begin{bmatrix} . & . & \cdots & . & . & \cdots & . & . \\ \tilde{a}_0. & \cdots & \tilde{a}_{n-k-2} & \tilde{a}_{n-k-1} & . & \cdots & \tilde{a}_{n-2} & \tilde{a}_{n-1} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & a_{n-k,n-1} \\ 0 & \cdots & . & 0 & 1 & \cdots & 0 & a_{n-k+1,n-1} \\ . & \cdots & . & . & . & \cdots & . & . \\ 0 & \cdots & & 0 & \cdots & 1 & 0 & a_{n-2,n-1} \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & a_{n-1,n-1} \end{bmatrix} =$$

$$[\tilde{a}_0, \tilde{a}_1 \dots, \tilde{a}_{n-k-2}, \tilde{a}_{n-k-1} + a_{n-1,n-1}, \tilde{a}_{n-k} + a_{n-1,n-1}^{(2)}, \dots, \tilde{a}_{n-2} + a_{n-1,n-1}^{(k)}, a_{n-1,n-1}^{(k+1)}].$$

The assumption (2.6) implies

$$e_{n-1}^T A^{k+1} = [0, \dots, 0, 1, a_{n-1,n-1}, a_{n-1,n-1}^{(2)}, \dots, a_{n-1,n-1}^{(k)}, a_{n-1,n-1}^{(k+1)}].$$

Hence $[\tilde{a}_0, \dots, \tilde{a}_{n-k-3}] = [0, \dots, 0]$, and $\tilde{a}_{n-k-2} = 1$. Moreover

$$[\tilde{a}_{n-k-1} + a_{n-1,n-1}, \tilde{a}_{n-k} + a_{n-1,n-1}^{(2)}, \dots, \tilde{a}_{n-2} + a_{n-1,n-1}^{(k)}] = [a_{n-1,n-1}, a_{n-1,n-1}^{(2)}, \dots, a_{n-1,n-1}^{(k)}]$$

yields $[\tilde{a}_{n-k-1}, \tilde{a}_{n-k}, \dots, \tilde{a}_{n-2}] = [0, \dots, 0]$. This proves (2.9) in the case $\nu = k+1$, and we obtain $A = H_n = F_a$ with $a = [a_{0,n-1}, \dots, a_{n-1,n-1}]^T$. □

It was proved in [4] that the map $(g, b) \mapsto u(A; g, b)$ is symmetric if $A = F_p$. The proof of [4, Proposition A.5] with the identities

$$e_{n-1}^T \left(\sum_{j=k}^{n-1} b_j A^{j-k} \right) g = e_{n-1}^T \left(\sum_{j=k}^{n-1} g_j A^{j-k} \right) b, \ \ k = 0, 1, \dots, n-1,$$

is rather involved. We remark that

$$L(F_p, e_{n-1}) = \begin{bmatrix} 1 & p_{n-1} & p_{n-2} & \cdots & p_1 \\ & 1 & p_{n-1} & \cdots & p_2 \\ & & \ddots & \ddots & \\ & & & . & . \\ & & & & 1 \end{bmatrix}.$$

The "crossover" identity (2.12) below appears in [4, Proposition A.3].

Theorem 2.4. *Suppose $A \in K^{n \times n}$ and*

$$\chi_A(z) = z^n - \sum\nolimits_{i=0}^{n-1} p_i z^i = z^n - p(z). \tag{2.11}$$

Let $g = [g_0, \dots, g_{n-1}]^T \in K^n$, $b = [b_0, \dots, b_{n-1}]^T \in K^n$, and $g_n, b_n \in K$. The following statements are equivalent.

(i) *The identity*

$$\sum\nolimits_{k=0}^{n} A^k g_k (b + b_n p) = \sum\nolimits_{k=0}^{n} A^k b_k (g + g_n p) \tag{2.12}$$

holds for all b, b_n, g, g_n.

(ii) *The matrix A is a second companion matrix such that $A = F_p$.*

Proof. (i) ⇒ (ii): Note that (2.12) is equivalent to

$$R(A, b)g + b_n R(A, p)g + g_n A^n b = R(A, g)b + g_n R(A, p)b + b_n A^n g. \tag{2.13}$$

If we choose $b_n = g_n = 0$ in (2.13) then we obtain $R(A, b)g = R(A, g)b$. Hence $A = F_q$ for some $q \in K^n$ (by Theorem 2.2). Then (2.11) implies $q = p$.

(ii) ⇒ (i): Assume $A = F_p$. Then $R(A, b)g = R(A, g)b$ for all $b, g \in K^n$. We also have $A^n = p(A)$ and therefore $A^n g = p(A)g = R(A, p)g$. Hence (2.13) holds for all b, g, b_n, g_n. □

3 Block companion matrices

In this section we extend results of the preceding section to block companion matrices. Let $P_i \in K^{t\times t}$, $i = 0,\dots,n-1$, be the entries of the block matrix

$$P = \begin{bmatrix} P_0 \\ \vdots \\ P_{n-1} \end{bmatrix}_{nt\times t},$$

and the coefficients of the monic matrix polynomial

$$\tilde{P}(z) = z^n I_t - (z^{n-1}P_{n-1} + \cdots + P_0).$$

A linearization [7] of $\tilde{P}(z)$ gives rise to the *block companion matrix of the second type*

$$F_P = \begin{bmatrix} 0 & \cdot & & & & P_0 \\ I_t & 0 & \cdot & & & P_1 \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ 0 & 0 & & & I_t & P_{n-1} \end{bmatrix} \in K^{nt\times nt}.$$

Our main tool is the Kronecker product. Let $S = (s_{ij}) \in K^{\ell\times m}$ and $T \in K^{p\times r}$ then $S \otimes T = (s_{ij}T) \in K^{\ell p\times mr}$. Define

$$E_0 = e_0 \otimes I_t = [I_t, 0, \dots, 0]^T, \dots, E_{n-1} = e_{n-1} \otimes I_t = [0, \dots, 0, I_t]^T.$$

To the block matrix

$$B = \begin{bmatrix} B_0 \\ B_1 \\ \vdots \\ B_{n-1} \end{bmatrix} = \sum_{i=0}^{n-1} E_i B_i \in K^{nt\times m}$$

with $B_j \in K^{t\times m}$, $j = 0,\dots,n-1$, we associate the matrix polynomial

$$B(z) = B_0 + zB_1 + \cdots + z^{n-1}B_{n-1} \in K^{t\times m}[z].$$

Let $A \in K^{nt\times nt}$. We define

$$B(A) = (I_n \otimes B_0) + A(I_n \otimes B_1) + \cdots + A^{n-1}(I_n \otimes B_{n-1}) \in K^{nt\times nm}. \tag{3.1}$$

In (3.1) we have an example of an operator substitution (see [8]). The matrix

$$R(A,B) = [B, AB, \dots, A^{n-1}B] \in K^{nt\times nm}$$

is the n-step reachability matrix of the pair (A,B). Let $B_i, G_i \in K^{t\times t}$, $i = 0,\dots,n-1$. We say that the $nt \times t$ matrices

$$B = \begin{bmatrix} B_0 \\ B_1 \\ \vdots \\ B_{n-1} \end{bmatrix} \quad \text{and} \quad G = \begin{bmatrix} G_0 \\ G_1 \\ \vdots \\ G_{n-1} \end{bmatrix}$$

are *blockwise commuting* if $B_iG_j = G_jB_i$, $i,j = 0,\dots,n-1$. Adapting arguments of the preceding section to Kronecker products we prove the following result.

Theorem 3.1. *Let $A \in K^{nt\times nt}$. The following statements are equivalent.*

(i) $A = F_P$ *for some* $P \in K^{nt\times t}$.
(ii) $R(A, E_0) = I_{nt}$.
(iii) $R(A, G) = G(A)$ *for all* $G \in K^{nt\times m}$.
(iv) *The identity*

$$[B, AB, \dots, A^{n-1}B]G = [G, AG, \dots, A^{n-1}G]B \tag{3.2}$$

holds for all blockwise commuting matrices $B, G \in K^{nt\times t}$.

Proof. We proceed along the lines suggested by a referee.

(ii) ⇒ (i): The implication is obvious.
(i) ⇒ (iii): For $A \in K^{nt\times nt}$ we have

$$R(A, G) = \sum_{i=0}^{n-1} R(A, E_i)(I_n \otimes G_i).$$

If $A = F_P$ then $R(A, E_i) = F_P^i$, and we obtain $R(A, G) = G(A)$.
(iii) ⇒ (iv): If B and G are blockwise commuting then

$$B(A)G = \sum_{i,j=0}^{n-1} A^i(I_n \otimes G_i)(I_n \otimes B_j) = G(A)B.$$

Thus (iii) implies $R(A, B)G = R(A, G)B$, which is the target identity (3.2).
(iv) ⇒ (ii): Choose $G = E_0$ in (3.2). Then it follows from $R(A, B)E_0 = B$ that $R(A, E_0)B = B$ holds for all $B \in K^{nt\times t}$. Hence $R(A, E_0) = I_{nt}$.

□

Acknowledgement: We are grateful to a referee for detailed comments and valuable suggestions, which helped us to improve the paper significantly.

References

[1] H. Bart and G. Ph. A. Thijsse, Simultaneous reduction to companion and triangular forms of sets of matrices, Linear Multilinear Algebra 26, 231–241 (1990).
[2] A. J. Brzezinski, Output-Only Techniques for Fault Detection, Dissertation, Department of Aerospace Engineering, University of Michigan, Ann Arbor, 2011.
[3] A. J. Brzezinski, E. Wu, and D. S. Bernstein, Curiously commuting vectors, Problem 44-3, IMAGE (Bulletin of the International Linear Algebra Society) 44 (2010).
[4] A. J. Brzezinski, S. Kukreja, Jun Ni, and D. S. Bernstein, Sensor-only fault detection using pseudo transfer function identification, in Proc. Amer. Contr. Conf., 5433–5438, Baltimore, June 2010.
[5] F. De Terán, F. M. Dopico, and J. Pérez, Condition numbers for inversion of Fiedler matrices, Linear Algebra Appl. 439, 944–981 (2013).
[6] A. Ferrante and H. K. Wimmer, Reachability matrices and cyclic matrices, Electron. J. Linear Algebra 20, 95–102 (2010).
[7] I. Gohberg, P. Lancaster, and L. Rodman, Matrix Polynomials, Academic Press, New York, 1982.
[8] M. L. J. Hautus, Operator substitution, Linear Algebra Appl. 205–206, 713–739 (1994).
[9] Th. Kailath, Linear Systems, Prentice Hall, Englewood Cliffs, 1980.