Extremal sparsity of the companion matrix of a polynomial*

Chao Ma, Xingzhi Zhan†
Department of Mathematics
East China Normal University
Shanghai 200241, China

Abstract

Let C be the companion matrix of a monic polynomial p over a field F. We prove that if A is a matrix whose entries are rational functions of the coefficients of p over F and whose characteristic polynomial is p, then A has at least as many nonzero entries as C.

Key words. Polynomial, companion matrix, sparsity, spanning branching, transcendence degree.

Mathematics subject classifications. 15A18, 12E05, 05C50, 05C05

The *companion matrix* of a monic polynomial

$$p(x) = x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n$$

over a field is defined to be

$$C(p) = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_n \\ 1 & 0 & \cdots & 0 & -a_{n-1} \\ 0 & 1 & \cdots & 0 & -a_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_1 \end{bmatrix}.$$

*E-mail addresses: machao0923@163.com(C.Ma), zhan@math.ecnu.edu.cn(X.Zhan). This research was supported by the NSFC grant 10971070.
†Corresponding author.

It is well known [5, p.147] that the characteristic polynomial of $C(p)$ is $p(x)$. Because of this relation, companion matrices can be used to study properties of polynomials. For example, matrix tools are used in [7] to locate the roots of a complex polynomial via its companion matrix.

The companion matrix $C(p)$ is very sparse, i.e., it has many zero entries. If we regard the coefficients $a_1, \ldots, a_n$ of $p(x)$ as distinct indeterminates, then $C(p)$ has $2n-1$ nonzero entries. We will show that the companion matrix is the sparsest in a sense to be described below.

Let F be a field and $x_1, \ldots, x_n$ be distinct indeterminates. We denote by $F[x_1, \ldots, x_n]$ the ring of polynomials in $x_1, \ldots, x_n$ over F, and by $F(x_1, \ldots, x_n)$ the field of rational functions in $x_1, \ldots, x_n$ over F :

$$F(x_1, \ldots, x_n) = \left\{ \frac{f}{g} \middle| f, g \in F[x_1, \ldots, x_n],\ g \neq 0 \right\}.$$

Denote by $M_n(E)$ the set of $n \times n$ matrices whose entries are elements of a given field E. Our main result is the following theorem.

Theorem 1. *Let F be a field, $a_1, \ldots, a_n$ be distinct indeterminates, and $A \in M_n(F(a_1, \ldots, a_n))$. If the characteristic polynomial of A is*

$$x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n, \tag{1}$$

then A has at least $2n-1$ nonzero entries.

To prove Theorem 1 we need several lemmas.

Lemma 2. *The polynomial in (1) is irreducible over $F(a_1, \ldots, a_n)$.*

Proof. Let $p(x) = x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n$. If $n = 1$, $p(x)$ is obviously irreducible. Now suppose $n \geq 2$, and we first show that $p(x)$ has no factors of degree 1. Otherwise $p(x)$ has a root f/g where $f, g \in F[a_1, \ldots, a_n]$ with $g \neq 0$. Then $f^n/g^n + a_1 f^{n-1}/g^{n-1} + \cdots + a_{n-1} f/g + a_n = 0$. Hence

$$f^n + a_1 f^{n-1} g + \cdots + a_{n-1} f g^{n-1} + a_n g^n = 0. \tag{2}$$

Note that $f \neq 0$. For a nonzero polynomial $h \in F[a_1, \ldots, a_n]$ we use $\deg_{a_n} h$ to denote the degree of a_n in h. Let $\deg_{a_n} f = d$ and $\deg_{a_n} g = e$, and let u be the polynomial on the left-hand side of (2). If $d \leq e$ then

$$\deg_{a_n} u = \deg_{a_n}(a_n g^n) = ne + 1 \geq 1,$$

contradicting (2); if $d > e$ then

$$\deg_{a_n} u = \deg_{a_n}(f^n) = nd \geq 2,$$

contradicting (2). Thus, $p(x)$ has no factors of degree 1. It follows that the lemma holds also for $n = 2$ or $n = 3$.

We use induction on n. Next suppose $n \geq 4$ and assume that the lemma holds for the degree $n-1$. To the contrary, suppose that $p(x)$ is reducible:

$$x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n = \left(\sum_{i=0}^{k} b_i x^{k-i}\right)\left(\sum_{j=0}^{n-k} c_j x^{n-k-j}\right) \tag{3}$$

where $b_i, c_j \in F(a_1, \ldots, a_n)$ with $b_0 = c_0 = 1$. Since we have proved that $p(x)$ has no factors of degree 1, $2 \leq k \leq n-2$. Given $h \in F[a_1, \ldots, a_n]$, we may write h as $h = h_0 + h_1 a_n + h_2 a_n^2 + \cdots + h_q a_n^q$ with each $h_i \in F[a_1, \ldots, a_{n-1}]$. Thus, if $h \neq 0$, $h = a_n^m w$ for some nonnegative integer m and $w \in F[a_1, \ldots, a_n]$ with $w(a_1, \ldots, a_{n-1}, 0) \neq 0$. Hence every nonzero $v \in F(a_1, \ldots, a_n)$ can be written as $v = v_1 / a_n^r$ where r is an integer, $v_1 \in F(a_1, \ldots, a_n)$ and $v_1(a_1, \ldots, a_{n-1}, 0) \neq 0$. Now let

$$b_i = \frac{f_i}{a_n^{s_i}}, \quad c_j = \frac{g_j}{a_n^{t_j}}$$

where $f_i, g_j \in F(a_1, \ldots, a_n)$, s_i and t_j are integers such that if $b_i \neq 0$ then $f_i(a_1, \ldots, a_{n-1}, 0) \neq 0$ and if $b_i = 0$ then $f_i = 0$ and $s_i = -1$, and if $c_j \neq 0$ then $g_j(a_1, \ldots, a_{n-1}, 0) \neq 0$ and if $c_j = 0$ then $g_j = 0$ and $t_j = -1$. Since $b_0 = c_0 = 1$, we set $f_0 = g_0 = 1$ and $s_0 = t_0 = 0$.

We will show that there are no positive integers among $s_0, \ldots, s_k, t_0, \ldots, t_{n-k}$. To the contrary we first suppose that there is at least one positive integer among $s_0, \ldots, s_k$. Let i_0 be the largest subscript such that $s_{i_0} = \max\{s_0, \ldots, s_k\}$. Then $s_{i_0} \geq 1$, $b_{i_0} \neq 0$ and $f_{i_0}(a_1, \ldots, a_{n-1}, 0) \neq 0$. We distinguish two cases.

Case 1. There is at least one nonnegative integer among $t_1, \ldots, t_{n-k}$. Let j_0 be the largest subscript such that $t_{j_0} = \max\{t_0, \ldots, t_{n-k}\}$. Then $t_{j_0} \geq 0$, $c_{j_0} \neq 0$ and $g_{j_0}(a_1, \ldots, a_{n-1}, 0) \neq 0$. Comparing the coefficients of $x^{n-i_0-j_0}$ on both sides of (3) we have

$$a_{i_0+j_0} = \sum_{i+j=i_0+j_0} b_i c_j = \sum_{i+j=i_0+j_0} \frac{f_i g_j}{a_n^{s_i+t_j}}. \tag{4}$$

Note that $s_{i_0} + t_{j_0} \geq s_i + t_j$ for all i, j with $i + j = i_0 + j_0$ and equality holds if and only if $i = i_0$ and $j = j_0$. Multiplying both sides of (4) by $a_n^{s_{i_0}+t_{j_0}}$ and then setting $a_n = 0$ we obtain

$$0 = f_{i_0}(a_1, \ldots, a_{n-1}, 0) g_{j_0}(a_1, \ldots, a_{n-1}, 0),$$

a contradiction.

Case 2. $t_1, \ldots, t_{n-k}$ are all negative integers. Comparing the coefficients of x^{n-i_0} on both sides of (3) we have

$$a_{i_0} = \sum_{i+j=i_0} b_i c_j = \sum_{i+j=i_0} \frac{f_i g_j}{a_n^{s_i+t_j}}. \tag{5}$$

Note that $s_{i_0} > s_i + t_j$ for all $i = 0, 1, \ldots, k$ and $j = 1, \ldots, n-k$. Multiplying both sides of (5) by $a_n^{s_{i_0}}$ and then setting $a_n = 0$ we obtain $0 = f_{i_0}(a_1, \ldots, a_{n-1}, 0)$, a contradiction.

Thus we have proved that $s_0, \ldots, s_k$ are all non-positive integers. In the same way we can prove that $t_0, \ldots, t_{n-k}$ are all non-positive integers. Consequently, all the $b_i(a_1, \ldots, a_{n-1}, 0)$ and $c_j(a_1, \ldots, a_{n-1}, 0)$ are well defined and they are elements of $F(a_1, \ldots, a_{n-1})$. Denote $\tilde{b}_i = b_i(a_1, \ldots, a_{n-1}, 0)$ and $\tilde{c}_j = c_j(a_1, \ldots, a_{n-1}, 0)$. Setting $a_n = 0$ in (3) we have

$$x^n + a_1 x^{n-1} + \cdots + a_{n-1} x = \left(\sum_{i=0}^{k} \tilde{b}_i x^{k-i} \right) \left(\sum_{j=0}^{n-k} \tilde{c}_j x^{n-k-j} \right). \tag{6}$$

Considering the constant term we get $\tilde{b}_k \tilde{c}_{n-k} = 0$. Hence $\tilde{b}_k = 0$ or $\tilde{c}_{n-k} = 0$. In either case (6) shows that $x^{n-1} + a_1 x^{n-2} + \cdots + a_{n-2} x + a_{n-1}$ is reducible over $F(a_1, \ldots, a_{n-1})$, which contradicts the induction hypothesis. This proves that $p(x)$ is irreducible.□

We will use a little graph theory [1, 9]. A *branching* is an oriented tree having a root of in-degree 0 and all other vertices of in-degree 1. A *spanning branching* of a digraph is a branching that includes all vertices of the digraph. If (a, b) is an arc of a digraph, then a is called an *in-neighbor* of b and b is called an *out-neighbor* of a. The following lemma is well known ([1, p.108] or [9, p.90]) and easy to prove.

Lemma 3. *In a strongly connected digraph, every vertex is the root of a spanning branching.*

We denote by $D(A)$ the digraph of a matrix $A = (a_{ij})$ of order n. The vertices of $D(A)$ are $1, 2, \ldots, n$ and (s, t) is an arc if and only if $a_{st} \neq 0$. $A(i, j)$ will mean the entry of the matrix A in row i and column j.

Lemma 4. *Let E be a field. If the digraph of a matrix $A \in M_n(E)$ has a spanning branching whose arcs are $(i_1, j_1), \ldots, (i_{n-1}, j_{n-1})$, then there exists a nonsingular diagonal matrix $G \in M_n(E)$ such that $GAG^{-1}(i_k, j_k) = 1$ for $k = 1, \ldots, n-1$.*

Proof. Let B be the spanning branching of $D(A)$. We renumber the vertices of B as follows. The root of B is 1. If 1 has t out-neighbors, number them as $2, 3, \ldots, t+1$

in any order. Then number the out-neighbors of $2, 3, \ldots, t+1$ successively. Continuing in this way we will finally renumber all the vertices of $D(A)$, since B is a spanning branching. Thus we obtain a new digraph D' with a spanning branching B'. The arcs of B' are $(r_1, 2), (r_2, 3), \ldots, (r_{n-1}, n)$ which satisfy the key condition $1 \leq r_k \leq k$ for each $k = 1, 2, \ldots, n-1$. In particular, $r_1 = 1$. Moreover, there is a permutation matrix P such that $D' = D(PAP^T)$ where P^T denotes the transpose of P. Let $PAP^T = (a'_{ij})$. Then $a'_{r_k, k+1} \neq 0$ for $k = 1, \ldots, n-1$.

We define n nonzero elements $d_1, \ldots, d_n$ successively by setting $d_1 = 1$, and $d_{k+1} = d_{r_k} a'_{r_k, k+1}$ for $k = 1, \ldots, n-1$. This is well defined since $1 \leq r_k \leq k$. Let $H = \text{diag}(d_1, d_2, \ldots, d_n)$. We have $HPAP^T H^{-1}(r_k, k+1) = 1$ for $k = 1, \ldots, n-1$. Let $G = P^T HP$. Then G is a nonsingular diagonal matrix and $GAG^{-1}(i_k, j_k) = 1$ for $k = 1, \ldots, n-1$. $\square$

Results similar to Lemma 4 are known ([3, p.259] and [8, p.3]).

We will use a little algebra [6]. Let $F \subseteq E$ be a field extension. We denote by $\text{trd}(E/F)$ the transcendence degree of E over F. Let $a_1, \ldots, a_n$ be distinct indeterminates. Since $\{a_1, \ldots, a_n\}$ is a transcendence basis of $F(a_1, \ldots, a_n)$ over F [6, p.317], we have $\text{trd}(F(a_1, \ldots, a_n)/F) = n$.

Given $e_1, \ldots, e_k \in E$, we denote by $F(e_1, \ldots, e_k)$ the subfield of E defined by

$$F(e_1, \ldots, e_k) = \left\{ \frac{f(e_1, \ldots, e_k)}{g(e_1, \ldots, e_k)} \,\middle|\, f, g \in F[x_1, \ldots, x_k],\ g(e_1, \ldots, e_k) \neq 0 \right\}.$$

It is easy to show that $\text{trd}(F(e_1, \ldots, e_k)/F) \leq k$.

Proof of Theorem 1. Recall that a square matrix R is said to be *reducible* if there exists a permutation matrix P such that PRP^T is of the form

$$\begin{bmatrix} R_1 & 0 \\ R_2 & R_3 \end{bmatrix}$$

where R_1 and R_3 are square matrices of order at least 1, i.e., they do appear, and 0 is a zero matrix. A square matrix that is not reducible is called *irreducible.* Obviously, the characteristic polynomial of a reducible matrix is reducible. Now by Lemma 2, the matrix A is irreducible. Since the digraph of an irreducible matrix is strongly connected [4, p.55], $D(A)$ is strongly connected. By Lemma 3, $D(A)$ has a spanning branching. Then by Lemma 4, there exists a nonsingular diagonal matrix $G \in M_n(F(a_1, \ldots, a_n))$ such that $A' = GAG^{-1}$ has at least $n-1$ entries equal to 1.

Suppose that A' has precisely m nonzero entries. Note that A' and A have the same characteristic polynomial in (1) and have the same number m of nonzero entries. Let the nonzero entries of A' be $e_1, e_2, \ldots, e_{m-n+1}, 1, \ldots, 1$. Of course, every $e_j \in F(a_1, \ldots, a_n)$ and hence $F(e_1, \ldots, e_{m-n+1}) \subseteq F(a_1, \ldots, a_n)$. On the other hand, since each of the coefficients $a_1, \ldots, a_n$ of the polynomial in (1) is the value of a polynomial over F in the entries of A', we have $F(a_1, \ldots, a_n) \subseteq F(e_1, \ldots, e_{m-n+1})$. Hence $F(e_1, \ldots, e_{m-n+1}) = F(a_1, \ldots, a_n)$. Finally from

$$n = \operatorname{trd}(F(a_1, \ldots, a_n)/F) = \operatorname{trd}(F(e_1, \ldots, e_{m-n+1})/F) \leq m - n + 1$$

we obtain $m \geq 2n - 1$. □

In the proof of Theorem 1 we have used a method in [2] and [3].

Acknowledgment. The authors are grateful to Dr. Lingfei Jin for helpful discussions.

References

[1] J.A. Bondy and U.S.R. Murty, Graph Theory, GTM 244, Springer, 2008.

[2] T. Britz, The inverse of a non-singular free matrix, Linear Algebra Appl., 338(2001), 245-249.

[3] T. Britz, J.J. McDonald, D.D. Olesky and P. van den Driessche, Minimal spectrally arbitrary sign patterns, SIAM J. Matrix Anal. Appl., 26(2004), no.1, 257-271.

[4] R.A. Brualdi and H.J. Ryser, Combinatorial Matrix Theory, Cambridge University Press, 1991.

[5] R.A. Horn and C.R. Johnson, Matrix Analysis, Cambridge University Press, 1985.

[6] T.W. Hungerford, Algebra, GTM 73, Springer, 1980.

[7] F. Kittaneh, Singular values of companion matrices and bounds on zeros of polynomials, SIAM J. Matrix Anal. Appl., 16(1995), no.1, 333-340.

[8] B.L. Shader, Notes on the $2n$-conjecture, AIM Workshop, 2006, Available from: http://www.aimath.org/pastworkshops/matrixspectrum.html

[9] D.B. West, Introduction to Graph Theory, Prentice Hall, NJ, 1996.