Acta Universitatis Carolinae. Mathematica et Physica

Ladislav Beran

A direct transformation between two fundamental matrix canonical forms

Acta Universitatis Carolinae. Mathematica et Physica, Vol. 27 (1986), No. 1, 41--47

Persistent URL: `http://dml.cz/dmlcz/142563`

A Direct Transformation between two Fundamental Matrix Canonical Forms

LADISLAV BERAN

Department of Algebra, Charles University, Prague*)

Received 14 May 1985

In this paper we investigate a matrix equation representing a transformation from a Jordan form into the corresponding rational canonical form and vice versa.

V tomto článku se vyšetřuje maticová rovnice představující transformaci Jordanova tvaru v odpovídající racionální kanonický tvar a obráceně.

В статье изучается матричное уравнение представляющее трансформацию формы Жордана в соответвующую рациональную каноническую форму и обратно.

1. Preliminaries

Let K denote an algebraically closed field. We shall denote the ring of all $n \times n$ matrices with entries in a ring R by $R_{n,n}$. Let $p(x) = x^d + p_{d-1}x^{d-1} + \ldots + p_0 \in \in R[x]$ and let

$$C(p(x)) = \begin{pmatrix} 0 & 0 & \ldots & 0 & -p_0 \\ 1 & 0 & \ldots & 0 & -p_1 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & -p_{d-2} \\ 0 & 0 & \ldots & 1 & -p_{d-1} \end{pmatrix}$$

be the companion matrix of the polynomial $p(x)$. Denote by

$$J_n(h) = \begin{pmatrix} h & 0 & \ldots & 0 & 0 \\ 1 & h & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 & h \end{pmatrix}$$

the $n \times n$ Jordan block belonging to $h \in R$.

*) 186 00 Prague 8, Sokolovská 83, Czechoslovakia.

Consider now a fixed matrix M of $K_{n,n}$. Then M can be interpreted as the matrix of a linear transformation f defined on a vector space V over K. Let V be expressed as a direct sum $Z_1 \oplus Z_2 \oplus \ldots \oplus Z_s$ of non-zero subspaces which are cyclic with respect to f. Such a decomposition of V enables to obtain the rational canonical form of M. It follows that there exists a regular matrix $T \in K_{n,n}$ such that

$$T^{-1}MT = \begin{pmatrix} C_1 & 0 & \ldots & 0 \\ 0 & C_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & C_s \end{pmatrix} = C$$

where every block C_i is the companion matrix $C_i = C((x - h_i)^{s_i})$ and $m_{ji}(x) = (x - h_i)^{s_i}$ is the minimal polynomial determined by the restriction f_i of the transformation f to Z_i.

On the other hand, under the assumption made above, we can consider the Jordan canonical form of M. Hence, there is a regular matrix $S \in K_{n,n}$ such that

$$S^{-1}MS = \begin{pmatrix} J_1 & 0 & \ldots & 0 \\ 0 & J_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & J_s \end{pmatrix} = J$$

where $J_i = J_{s_i}(h_i)$ denotes the Jordan block belonging to h_i. (For the proofs of these remarks and associated discussion the reader is referred to [1].)

Consequently, from these general considerations we can easily conclude that there is a regular matrix P satisfying $P^{-1}JP = C$.

Evidently, if P_i denotes a matrix such that $P_i^{-1} J_i P_i = C_i$ then it suffices to put

$$P = \begin{pmatrix} P_1 & 0 & \ldots & 0 \\ 0 & P_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & P_s \end{pmatrix}.$$

From this point of view it is clear why the equation

$$X^{-1} J_n(h) X = C((x - h)^n) \tag{1}$$

may be worth investigating more closely.

The natural questions to ask about the equation are the following:

1. Is there a matrix satisfying (1) and having a nice and concretely describable form?

2. What can be said about all solutions of (1)?

The purpose of this note is to give an answer to these questions. In what follows we assume that R denotes an associative ring with identity element 1. Moreover, we suppose that $n \geqq 2$.

2. Particular solution

Convention. In this section R denotes a ring which need not be commutative.

We shall need the following lemma which is well-known. A proof is included for completeness.

2.1. Lemma. Let $0 \leqq p < n$. Then

$$\sum_{j=p}^{n}(-1)^j\binom{j}{p}\binom{n}{j} = 0\,.$$

Proof. Since

$$\binom{b}{a}\binom{b-a}{c-a} = \binom{c}{a}\binom{b}{c}$$

for any natural integers $a \leqq c \leqq b$, we have

$$\sum_{j=p}^{n}(-1)^j\binom{j}{p}\binom{n}{j} = \binom{n}{p}\sum_{j=p}^{n}(-1)^j\binom{n-p}{j-p} = \binom{n}{p}(-1)^p\sum_{s=0}^{n-p}(-1)^s\binom{n-p}{s} = 0\,.$$

The lemma thus follows.

Now let $h \in R$ and let

$P_n(h) =$

$$\begin{pmatrix}
1 & h & h^2 & \cdots & & \binom{k}{0}h^k & \cdots & \binom{n-1}{0}h^{n-1} \\
0 & 1 & 2h & \cdots & & \binom{k}{1}h^{k-1} & \cdots & \binom{n-1}{1}h^{n-2} \\
0 & 0 & 1 & \cdots & & \binom{k}{2}h^{k-2} & \cdots & \binom{n-1}{2}h^{n-3} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & 0 & \cdots\ 1, & \binom{k}{1}h, & \binom{k+1}{2}h, & \ldots, & \binom{n-1}{n-k}h^{n-k} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & & & \vdots & & & & \\
0 & 0 & 0 & \cdots\ 0 & & & \cdots & 1 \\
 & & & \uparrow & & & & \\
 & & & (k+1)\text{-st column} & & & &
\end{pmatrix} \begin{matrix} \\ \\ \\ \\ \leftarrow k\text{-th row} \\ \\ \\ \\ \\ \\ \end{matrix}$$

2.2. Proposition. Let h be an element of R. Then $P_n^{-1}(h) = P_n(-h)$.

Proof. Write $P_n(-h)\,P_n(h) = (c_{ki})$. It suffices to show that $c_{ki} = 0$ whenever

$k < i$. However, by Lemma 2.1, in this case

$$c_{ki} = 1.\binom{i-1}{k-1} h^{i-k} - \binom{k}{1} h.\binom{i-1}{k} h^{i-k-1} +$$

$$+ \binom{k+1}{2} h^2.\binom{i-1}{k+1} h^{i-k-2} - \ldots + \binom{i-1}{i-k} h^{i-k}(-1)^{i-k}.\binom{i-1}{i-1} =$$

$$= (-1)^{k-1} h^{i-k} \sum_{j=k-1}^{i-1} \binom{j}{k-1}\binom{i-1}{j}(-1)^j = 0.$$

This completes the proof of Proposition 2.2.

In the proof of the theorem 2.3 below we shall need the following explicit form of the companion matrix $C_n(h) = C((x-h)^n)$ associated to the polynomial $(x-h)^n$:

$$C_n(h) = \begin{pmatrix} 0\ 0\ 0\ \ldots\ 0 & (-1)^{n+1}\binom{n}{n} h^n \\ 1\ 0\ 0\ \ldots\ 0 & (-1)^n\binom{n}{n-1} h^{n-1} \\ 0\ 1\ 0\ \ldots\ 0 & (-1)^{n-1}\binom{n}{n-2} h^{n-2} \\ \cdots\cdots\cdots & \cdots\cdots\cdots \\ 0\ 0\ 0\ \ldots\ 0 & -\binom{n}{2} h^2 \\ 0\ 0\ 0\ \ldots\ 1 & \binom{n}{1} h \end{pmatrix}.$$

We are now in a position to state our main result of this section.

2.3. Theorem. Suppose h is an element of R. Then the matrix $P_n(h)$ is a particular solution of the equation (1).

Proof. Let $J_n(h)\,P_n(h) = (d_{st})$ and $P_n(h)\,C_n(h) = (e_{st})$. In order to prove the assertion, it suffices to show that $d_{st} = e_{st}$ for every $1 \leqq s \leqq t \leqq n$.

In the present context, two cases distinguish themselves:

Case I: $1 \leqq s \leqq t \leqq n - 1$. Here

$$e_{st} = \binom{s+u-1}{u} h^u$$

where $u = t - s + 1$. Hence, evidently, $e_{st} = d_{st}$.

Case II: $1 \leqq s \leqq t = n$. Then

$$e_{sn} = 1(-1)^{ns}\binom{n}{n-s+1} h^{n-s+1} + \binom{s}{1} h(-1)^{n-s+1}\binom{n}{n-s} h^{n-s} + \ldots$$

$$\dots + \binom{n-1}{n-s} h^{n-s} \binom{n}{1} h = h^{n-s+1}(-1)^{n+1} \sum_{j=s-1}^{n-1} \binom{j}{s-1}\binom{n}{j}(-1)^j .$$

Directly from Lemma 2.1 we next infer that

$$\sum_{j=s-1}^{n-1} \binom{j}{s-1}\binom{n}{j}(-1)^j = \sum_{j=s-1}^{n} (-1)^j \binom{j}{s-1}\binom{n}{j} - (-1)^n \binom{n}{s-1}\binom{n}{n} =$$

$$= (-1)^{n+1} \binom{n}{n-s+1} .$$

Therefore,

$$e_{sn} = h^{n-s+1}(-1)^{n+1}(-1)^{n+1} \binom{n}{n-s+1} = d_{sn}$$

and the theorem is proved.

3. General Solution

Convention. In this section we assume that R is commutative.

At this point we first need the following fact which becomes crucial in the sequel: the centralizer of a Jordan block in $R_{n,n}$ has a particularly simple structure. Indeed, we have

3.1. Proposition. For all $h \in R$ a matrix $B_n = (b_{ij}) \in R_{n,n}$ satisfies

$$B_n J_n(h) = J_n(h) B_n \tag{2}$$

if and only if it is of the form

$$B_n = \begin{pmatrix} b_{11} & 0 & 0 & \dots & 0 \\ b_{21} & b_{11} & 0 & \dots & 0 \\ b_{31} & b_{21} & b_{11} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ b_{n,1} & b_{n-1,1} & & \dots & b_{11} \end{pmatrix} \tag{3}$$

where $b_{11}, b_{21}, \dots, b_{n1}$ are arbitratily chosen in R. It is regular if and only if $b_{11}^n \neq 0$.

Proof. We shall use induction on n. However, a small trick is needed. Let $B_n\, J_n(h) = (f_{ij})$, $J_n(h)\, B_n = (g_{ij})$.

Suppose now that $(f_{ij}) = (g_{ij})$. Then $b_{11}h + b_{12} = f_{11} = g_{11} = b_{11}h$ and so $b_{12} = 0$. From $f_{22} = g_{22}, \dots, f_{n-1,n-1} = g_{n-1,n-1}$ we successively obtain $b_{23} = 0, \dots, b_{n-1,n} = 0$. Thus

$$b_{i,i+1} = 0 \tag{4}$$

for every $i = 1, 2, \dots, n-1$.

To begin the proof by induction, consider $n = 2$. A simple argument shows that in this case B_2 satisfies (2) if and only if it is of the form (3). Proceeding by induction, we assume the assertion true for $n - 1$ ($n \geqq 3$). Rewriting the left hand member of (2) and the right hand one and dividing the matrices B_n and $J_n(h)$ into blocks, we get

$$(5) \qquad B_n J_n(h) = \left(\begin{array}{c|c} B_{n-1} & \begin{matrix} b_{1n} \\ \vdots \\ b_{n-1,n} \end{matrix} \\ \hline b_{n1}, \ldots, b_{n,n-1} & b_{nn} \end{array}\right) \left(\begin{array}{c|c} J_{n-1}(h) & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline 0 \ \ldots \ 1 & h \end{array}\right) =$$

$$= \left(\begin{array}{c|c} B_{n-1} J_{n-1}(h) + \begin{pmatrix} 0 & \ldots & 0 & b_{1n} \\ 0 & \ldots & 0 & b_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & b_{n-1,n} \end{pmatrix} & \begin{matrix} b_{1n}h \\ b_{2n}h \\ \vdots \\ b_{n-1,n}h \end{matrix} \\ \hline b_{n1}h + b_{n2}, \ldots, b_{n,n-1}h + b_{nn} & b_{nn}h \end{array}\right);$$

$$(6) \qquad J_n(h) B_n = \left(\begin{array}{c|c} J_{n-1}(h) & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline 0 \ \ldots \ 1 & h \end{array}\right) \left(\begin{array}{c|c} B_{n-1} & \begin{matrix} b_{1n} \\ \vdots \\ b_{n-1,n} \end{matrix} \\ \hline b_{n1}, \ldots, b_{n,n-1} & b_{nn} \end{array}\right) =$$

$$= \left(\begin{array}{c|c} J_{n-1}(h) B_{n-1} & \begin{matrix} b_{1n}h \\ b_{1n} + b_{2n}h \\ \vdots \\ b_{n-2,n} + b_{n-1,n}h \end{matrix} \\ \hline b_{n-1,1} + b_{n1}h, \ldots, b_{n-1,n-1} + b_{n,n-1}h & b_{n-1,n} + b_{nn}h \end{array}\right).$$

From (5), (6) and (4) we can easily see that (2) implies

$$(7) \qquad J_{n-1}(h) B_{n-1} = B_{n-1} J_{n-1}(h)$$

as well as the conditions

$$(8) \qquad b_{1n} = 0, \ldots, b_{n-2,n} = 0 ;$$

$$(9) \qquad b_{n-1,1} = b_{n2}, \quad b_{n-1,2} = b_{n3}, \ldots, b_{n-1,n-1} = b_{nn} .$$

Hence, by applying the induction assumption, we conclude that B_n is of the form given in (3).

Conversely, if B_n is of the form (3), then, clearly, $b_{in} = 0$ for every $i = 1, 2, \ldots$ $\ldots, n - 1$. The proof of the equality (2) is then immediate upon noting that one can use a short inductive argument based on (5) and (6).

Thus we have now proved everything claimed in Proposition 3.2.

Combining Proposition 3.1 with Theorem 2.3, we have the following theorem.

3.2. Theorem. Let $h \in R$ and let $Q_n \in R_{n,n}$. Then $Q_n^{-1} J_n(h) Q_n = C((x - h)^n)$ if and only if $Q_n = B_n P_n(h)$ where $P_n(h)$ is the matrix described above and B_n is of the form specified in (3) with $b_{11}^n \neq 0$.

Proof. By Theorem 2.3, $Q_n^{-1} J_n(h) Q_n = C((x - h)^n) = C_n(h)$ is true if and only if $Q_n^{-1} J_n(h) Q_n = P_n^{-1}(h) J_n(h) P_n(h)$. Hence, the considered equality is equivalent to $Q_n P_n^{-1}(h) = B_n$ where $B_n^{-1} J_n(h) B_n = J_n$. The theorem is now immediate from Proposition 3.1.

References

[1] Herstein, I. N.: Topics in Algebra, Blaisdell Publ. Co., London 1964.