A matrix equation and the Jordan canonical form*

Chi-Kwong Li
Department of Mathematics, College of William & Mary,
Williamsburg, VA 23187, USA.
ckli@math.wm.edu

Abstract

For a given $p \times q$ complex matrix C, a necessary and sufficient condition is obtained for the existence of a matrix X satisfying $J_pX - XJ_q = C$, here J_r denotes the $r \times r$ Jordan block of 0. An easy construction of the solution X is given if it exists. These results lead to a proof of the fact that a nilpotent matrix is similar to a direct sum of Jordan blocks.

AMS Classification. 15A20, 15A24.
Keywords. Jordan canonical form, nilpotent matrix, matrix equation.

1 Introduction

Let $M_{m,n}$ be the set of $m \times n$ complex matrices, and $M_n = M_{n,n}$. Denote by E_{ij} the standard matrix unit with the (i,j) entry equal to 1 and the other entries equal to 0. The size of E_{ij} should be clear in the context.

For $\lambda \in \mathbb{C}$ and a positive integer r, the matrix

$$J_r(\lambda) = \lambda I_r + \sum_{j=1}^{r-1} E_{j,j+1} = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ & & & \lambda \end{pmatrix} \in M_r$$

is called the Jordan block of λ of size r. We have the following Jordan canonical form theorem; e.g., see [1, Chapter 12] for a proof and some historical notes.

Theorem 1.1. *Every matrix $A \in M_n$ is similar to a direct sum of Jordan blocks.*

The result has many interesting consequences, and has applications to other topics; for example, see [1, Chapter 13]. The theorem can be proved by establishing the following two assertions.

Assertion 1 *A matrix $A \in M_n$ with distinct eigenvalues $\lambda_1, \dots, \lambda_k$, is similar to a direct sum of square matrices $A_1, \dots, A_k$, denoted by $A_1 \oplus \cdots \oplus A_k$, such that A_j has λ_j as the only (distinct) eigenvalue for $j = 1, \dots, k$.*

Assertion 2 *If $B \in M_n$ has only one distinct eigenvalue λ, then B is similar to a direct sum of Jordan blocks of λ.*

*Research partially supported by the Simons Foundation Grant 851334

By these assertions, there are invertible matrices $R, R_1, \ldots, R_k$ such that $R^{-1}AR = A_1 \oplus \cdots \oplus A_k$, and $R_j^{-1}A_jR_j$ is a direct sum of Jordan blocks of λ_j. Then $S^{-1}AS$ is a direct sum of Jordan blocks if $S = R(R_1 \oplus \cdots \oplus R_k)$.

One approach to prove Assertion 1 is to use the Sylvester equation theorem, which asserts that the matrix equation $FX - XG = C$ always has a unique solution $X \in M_{p,q}$ for given $F \in M_p, G \in M_q$ and $C \in M_{p,q}$ such that F and G have no common eigenvalues; see [2], [1, Theorem 11.4.1], and Lemma 2.3 in the next section. We will show that Assertion 2 can be proved by using a solution of the matrix equation $J_p(0)X - XJ_q(0) = C$ for a given matrix $C \in M_{p,q}$. In Theorem 2.1, a necessary and sufficient condition is obtained for the existence of a solution $X \in M_{p,q}$ of the equation. An easy construction of a solution X is given if it exists. The result will then be used to give a proof of Assertion 2. For completeness, we will also give a proof of Assertion 1 and some related remarks.

2 Auxiliary results and proofs

For positive integers r and s, we let $J_s = J_s(0)$, and note that $J_s^r = 0$ if and only if $r \geq s$.

Theorem 2.1. *Let p, q be positive integers with $p \geq q \geq 1$, let $C = (c_{ij}) \in M_{p,q}$, and let*

$$T = \begin{pmatrix} J_p & C \\ 0 & J_q \end{pmatrix}.$$

The following conditions are equivalent.

(a) *There is $X \in M_{p,q}$ such that $J_pX - XJ_q = C$.*

(b) *The matrix T is similar to $J_p \oplus J_q$.*

(c) *$T^p = 0$.*

(d) *$(c_{p,1}, \ldots, c_{p,q}) + (0, c_{p-1,1}, \ldots, c_{p-1,q-1}) + \cdots + (0, \ldots, 0, c_{p-q+1,1}) = (0, \ldots, 0)$.*

Moreover, if (d) *holds, and if $X = (x_{ij}) \in M_{p,q}$, in which $(x_{11}, \ldots, x_{1q}) = (0, \ldots, 0)$, and*

$$(x_{\ell,1}, \ldots, x_{\ell,q}) = (c_{\ell-1,1}, \ldots, c_{\ell-1,q}) + (0, x_{\ell-1,1}, \ldots, x_{\ell-1,q-1}), \quad \ell = 2, \ldots, p,$$

then $J_pX - XJ_q = C$ and

$$\begin{pmatrix} I_p & -X \\ 0 & I_q \end{pmatrix} \begin{pmatrix} J_p & C \\ 0 & J_q \end{pmatrix} \begin{pmatrix} I_p & X \\ 0 & I_q \end{pmatrix} = \begin{pmatrix} J_p & 0 \\ 0 & J_q \end{pmatrix}. \qquad (1)$$

By Theorem 2.1, one can use the simple condition (d) to determine whether condition (a), (b), or (c) holds. Moreover, if condition (d) holds, one can construct X satisfying (a). The matrix X will also satisfy (1), and hence condition (b) holds.

Proof of Theorem 2.1. Suppose (a) holds. Then the matrix X will satisfy (1) so that T is similar to $J_p \oplus J_q$. Thus, condition (b) holds.

Suppose (b) holds. Then there is an invertible matrix S such that $T = S^{-1}(J_p \oplus J_q)S$ so that $T^p = S^{-1}(J_p^p \oplus J_q^p)S = 0_{p+q}$ as $J_p^p = 0_p$ and $J_q^p = 0_q$. Thus, condition (c) holds.

Suppose (c) holds. By an easy induction argument, one can show that

$$T^\ell = \begin{pmatrix} J_p^\ell & Q_\ell \\ 0 & J_q^\ell \end{pmatrix} \text{ with } Q_\ell = J_p^{\ell-1}C + J_p^{\ell-2}CJ_q + \cdots + J_pCJ_q^{\ell-2} + CJ_q^{\ell-1}, \quad \ell = 2, 3, \ldots.$$

If $C = (c_{ij}) \in M_{p,q}$ has rows $C_1, \ldots, C_p$, then $T^p = \begin{pmatrix} 0_p & Q_p \\ 0 & 0_q \end{pmatrix}$ with

$$Q_p = \begin{pmatrix} C_p \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} C_{p-1} \\ C_p \\ 0 \\ \vdots \\ 0 \end{pmatrix} J_q + \cdots + \begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ \vdots \\ C_p \end{pmatrix} J_q^{p-1}.$$

Let Q_p have rows $Y_1, Y_2, \ldots, Y_q$. Then

$$Y_1 = C_p + C_{p-1}J_q + \cdots + C_1J_q^{p-1} = (c_{p,1}, c_{p,2} + c_{p-1,1}, \ldots, c_{p,q} + \cdots + c_{p-q+1,q}),$$

and for $\ell > 1$,

$$Y_\ell = C_pJ_q^{\ell-1} + \cdots + C_\ell J_q^{p-1} = (C_p + \cdots + C_1J_q^{p-1})J_1^{\ell-1} = Y_1J_q^{\ell-1},$$

here we use the fact that $J_p^j = 0$ for $j \geq p$ to get the second equality. Thus, $T^p = 0$ if and only if

$$0 = Y_1 = C_p + C_{p-1}J_q + \cdots + C_1J_q^{p-1},$$

which is the vector on the left hand side in (d).

Finally, suppose (d) holds. Let X be defined as in the last assertion of the theorem. If $Z = (z_{ij}) = J_pX - XJ_q$, then

$$Z = \begin{pmatrix} x_{2,1} & x_{2,2} & \cdots & x_{2,q-1} & x_{2,q} \\ x_{3,1} & x_{3,2} & \cdots & x_{3,q-1} & x_{3,q} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{p,1} & x_{p,2} & \cdots & x_{p,q-1} & x_{p,q} \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & x_{1,1} & x_{1,2} & \cdots & x_{1,q-1} \\ 0 & x_{2,1} & x_{2,2} & \cdots & x_{2,q-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & x_{p-1,1} & x_{p-1,2} & \cdots & x_{p-1,q-1} \\ 0 & x_{p,1} & x_{p,2} & \cdots & x_{p,q-1} \end{pmatrix}.$$

By the definition of X, for $\ell = 1, \ldots, p-1$, the ℓth row of Z equals

$$(x_{\ell+1,1}, \ldots, x_{\ell+1,q}) - (0, x_{\ell,1}, \ldots, x_{\ell,q-1}) = (c_{\ell,1}, \ldots, c_{\ell,q}),$$

and the last row of Z equals

$$\begin{aligned} & -(0, x_{p,1}, \ldots, x_{p,q-1}) = -(0, c_{p-1,1}, \ldots, c_{p-1,q-1}) - (0, 0, x_{p-1,1}, \ldots, x_{p-1,q-2}) \\ = \; & -\{(0, c_{p-1,1}, \ldots, c_{p-1,q-1}) + (0, 0, c_{p-2,1}, \ldots, c_{p-2,q-2}) + (0, 0, 0, x_{p-2,1}, \ldots, x_{p-2,q-3})\} \\ = \; & \cdots = -\{(0, c_{p-1,1}, \ldots, c_{p-1,q-1}) + \cdots + (0, \ldots, 0, c_{p-q+1,1})\} = (c_{p,1}, \ldots, c_{p,q}) \end{aligned}$$

by condition (d). Thus, $J_pX - XJ_q = C$, i.e., condition (a) holds.

Now, X satisfies (a). It will also satisfy (1). So, the last assertion of the theorem holds. □

We can use Theorem 2.1 to give the following.

Proof of Assertion 2. Suppose $B \in M_n$ has only one (distinct) eigenvalue λ. We only need to show that for $T = B - \lambda I_n$ there is R such that $R^{-1}TR$ is a direct sum of Jordan blocks of 0. Then $R^{-1}BR$ is a direct sum of Jordan blocks of λ.

Note that $T^n = 0$. We can find the smallest integer p such that $T^{p-1} \neq 0$ and $T^p = 0$. Then there is $v \in \mathbb{C}^n$ such that $T^{k-1}v \neq 0$ and $T^p v = 0$. We will show that $\{v, Tv, \ldots, T^{p-1}v\}$ is a linearly independent set. Suppose $\sum_{j=0}^{p-1} \alpha_j T^j v = 0$ with some $\alpha_\ell \neq 0$. Let ℓ be the smallest nonnegative integer such that $\alpha_\ell \neq 0$. Then $T^\ell v = \sum_{j=\ell+1}^{p-1}(-\alpha_j/\alpha_\ell)T^j v$ and

$$T^{p-1}v = T^{p-1-\ell}(T^\ell v) = T^{p-1-\ell}\left(\sum_{j=\ell+1}^{p-1}(-\alpha_j/\alpha_\ell)T^j v\right) = 0,$$

which contradicts the assumption that $T^{p-1}v \neq 0$.

Let $R_1 \in M_n$ be invertible with $T^{p-1}v, T^{p-2}v, \ldots, v$ as its first p columns. If $p = n$, then $R_1^{-1}TR_1 = J_n$ and we are done. Otherwise, $R_1^{-1}TR_1 = \begin{pmatrix} T_{11} & T_{12} \\ 0 & T_{22}\end{pmatrix}$ such that $T_{11} = J_p$ and $T_{22}^p = 0$. By induction assumption, there is an invertible $R_2 \in M_{n-p}$ such that $R_2^{-1}T_{22}R_2 = J_{n_2} \oplus \cdots \oplus J_{n_k}$ with $n_2 \geq \cdots \geq n_k$. Thus, $G = (I_p \oplus R_2^{-1})R_1^{-1}TR_1(I_p \oplus R_2) = (G_{ij})$ such that $G_{11} = J_p$, $G_{jj} = J_{n_j}$ for $j = 2, \ldots, k$, and $G_{ij} = 0_{n_i,n_j}$ whenever $i \neq j$ and $i \neq 1$.

For $j = 2, \ldots, k$, let $F_j = \begin{pmatrix} J_p & G_{1j} \\ 0 & J_{n_j}\end{pmatrix}$. Then F_j^p is a principal submatrix of $G^p = 0$. So, $F_j^p = 0$. By Theorem 2.1, for each $j = 2, \ldots, k$, there is $X_j \in M_{p,n_j}$ such that

$$\begin{pmatrix} J_p & G_{1j} \\ 0 & J_{n_j}\end{pmatrix}\begin{pmatrix} I_p & X_j \\ 0 & I_{n_j}\end{pmatrix} = \begin{pmatrix} I_p & X_j \\ 0 & I_{n_j}\end{pmatrix}\begin{pmatrix} J_p & 0 \\ 0 & J_{n_j}\end{pmatrix}.$$

Let $R_3 = \begin{pmatrix} I_p & X \\ 0 & I_{n-p}\end{pmatrix}$ with $X = (X_2 \cdots X_k)$. Then $(G_{ij})R_3 = R_3(J_p \oplus J_{n_2} \oplus \cdots \oplus J_{n_k})$. Let $R = R_1(I_p \oplus R_2)R_3$. Then $R^{-1}TR$ is a direct sum of Jordan blocks, and so is $R^{-1}BR$. □

For completeness, we also present a proof of Assertion 1 and some related remarks. In particular, one may see how Lemma 2.3 motivates the formulation of Theorem 2.1.

Lemma 2.2. *Suppose $A \in M_n$ has eigenvalues $\lambda_1, \ldots, \lambda_n$. There is an invertible $R \in M_n$ such that $R^{-1}AR$ is in upper triangular form with diagonal entries $\lambda_1, \ldots, \lambda_n$.*

Proof. We prove the result by the induction on n. The result is trivial if $n = 1$. Assume $n > 1$, and the result holds for matrices in M_{n-1}. Suppose $Ax = \lambda_1 x$ for a nonzero vector x. Let $R_1 \in M_n$ be invertible with its first column equal to x. Then $R_1^{-1}AR_1 = \begin{pmatrix} \lambda_1 & \star \\ 0 & A_1\end{pmatrix}$. Since $\det(xI - A) = (x - \lambda_1)\det(xI - A_1)$, we see that A_2 has eigenvalues $\lambda_2, \ldots, \lambda_n$. By induction assumption, there is an invertible $R_2 \in M_{n-1}$ such that $R_2^{-1}A_1R_2$ is in upper triangular form with diagonal entries $\lambda_2, \ldots, \lambda_n$. Let $R = R_1([1] \oplus R_2)$. Then $R^{-1}AR$ is in upper triangular form with diagonal entries $\lambda_1, \ldots, \lambda_n$. □

The matrices R_1 and R_2 in the proof can be chosen to be unitary if we use the inner product structure of $\mathbb{C}^n$. One can then conclude that for every $A \in M_n$ there is a unitary matrix $U \in M_n$

such that U^*AU is in upper triangular form. This is known as the Schur triangularization lemma; e.g., see [1, Theorem 11.1.1].

Lemma 2.3. *Suppose $F \in M_p, G \in M_q$ have no common eigenvalues, and $C \in M_{p,q}$. There is a unique matrix $X \in M_{p,q}$ such that $FX+C = XG$. As a result, if $R = \begin{pmatrix} I_p & X \\ 0 & I_q \end{pmatrix}$ and $A = \begin{pmatrix} F & C \\ 0 & G \end{pmatrix}$, then $R^{-1}AR = F \oplus G$.*

Proof. Let $R \in M_q$ be invertible such that $\tilde{G} = R^{-1}GR$ is in upper triangular form. Suppose $\tilde{C} = CR$ and $Y = XR$. Then $FX + C = XG$ if and only if $FY + \tilde{C} = Y\tilde{G}$. We will show that the modified equation $\tilde{C} = -FY + Y\tilde{G}$ has a unique solution Y. Then $X = YR^{-1}$ will be the unique solution of the original equation. One can check that $AR = R(F \oplus G)$ so that the last assertion of the lemma follows.

Let $\tilde{C} = (c_1 \cdots c_q)$ and $Y = (y_1 \cdots y_q)$ with $c_1, \ldots, c_q, y_1, \ldots, y_q \in \mathbb{C}^p$. If $\tilde{G} = (g_{ij})$, then $g_{11}, \ldots, g_{qq}$ are the eigenvalues of G. Then g_{jj} is not an eigenvalue of F so that $F - g_{jj}I_p$ is invertible for $j = 1, \ldots, q$. As a result, $Fy_1 + c_1 = g_{11}y_1$ has a unique solution $y_1 = -(F - g_{11}I_p)^{-1}c_1$, and for $\ell = 2, \ldots, q$,

$Fy_\ell + c_\ell = g_{\ell\ell}y_\ell + \sum_{j=1}^{\ell-1} g_{1j}y_\ell$ has a unique solution $y_\ell = (F - g_{\ell\ell}I_p)^{-1}(\sum_{j=1}^{\ell-1} g_{1j}y_j - c_\ell)$.

Thus, we get the unique solution $Y = [y_1 \cdots y_q]$ such that $FY + \tilde{C} = Y\tilde{G}$. □

Note that our proof of Lemma 2.3 provides an easy computational scheme for solving the Sylvester equation $FX - XG = C$. We can now present the following.

Proof of Assertion 1 We prove the result by the induction on k, the number of distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ of $A \in M_n$. If $k = 1$, the result is trivial. Assume that the result holds for matrices with fewer than k distinct eigenvalues for $k > 1$. Let $A \in M_n$ have k distinct eigenvalues. By Lemma 2.2, there is an invertible matrix $R_1 \in M_n$ such that $R_1^{-1}AR_1 = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$, where $A_{11} \in M_p$ is in upper triangular form with all diagonal entries equal to λ_1, and $A_{22} \in M_{n-p}$ is in upper triangular form with diagonal entries in $\{\lambda_2, \ldots, \lambda_k\}$. By Lemma 2.3, there is $X \in M_{p,n-p}$ such that $A_{11}X + A_{12} = XA_{22}$. Let $R_2 = \begin{pmatrix} I_p & X \\ 0 & I_{n-p} \end{pmatrix}$ so that $AR_2 = R_2 \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}$. By induction assumption, there is an invertible matrix $R_3 \in M_{n-p}$ such that $R_3^{-1}A_{22}R_3$ is a direct sum of diagonal blocks of matrices $A_2, \ldots, A_k$ such that each B_j is in triangular form with constant diagonal entry. Let $S = R_1R_2(I_p \oplus R_3)$. Then $R^{-1}AR$ has the desired form. □

Acknowledgment The author would like to thank the referee for some helpful comments.

Declaration There is no conflict of interest concerning this research.

References

[1] S.R. Garcia and R.A. Horn, Matrix Mathematics: A Second Course in Linear Algebra (Cambridge Mathematical Textbooks) 2nd Edition, Cambridge University Press, 2023.

[2] J. Sylvester, Sur l'equations en matrices $px = xq$, C. R. Acad. Sci. Paris. 99 (2) (1884), 67-71. 115–116.