

FACTORIZATION OF SINGULAR MATRICES

by

A.R. SOUROUR AND KUNIKYO TANG

DMS-565-IR

**November 1990
[Revised March 1991]**

FACTORIZATION OF SINGULAR MATRICES

A.R. SOUROUR AND KUNIKYO TANG

ABSTRACT. We give a necessary and sufficient condition that a singular square matrix A over an arbitrary field can be written as a product of two matrices with prescribed eigenvalues. Except when A is a 2×2 nonzero nilpotent, the condition is that the number of zeros among the eigenvalues of the factors is not less than the nullity of A . We use this result to prove results about products of hermitian and positive semidefinite matrices simplifying and strengthening some known results.

INTRODUCTION AND STATEMENT OF MAIN RESULT

In [5] the first author showed that if A is a nonscalar invertible $n \times n$ matrix over a field F and if $\beta_1, \beta_2, \dots, \beta_n, \gamma_1, \gamma_2, \dots, \gamma_n$ are elements of F , then A can be written as a product BC , where the eigenvalues of B and C are β_1, \dots, β_n and $\gamma_1, \dots, \gamma_n$ respectively, if and only if $\beta_1 \gamma_1 \cdots \beta_n \gamma_n = \det A$. In this article we give a necessary and sufficient condition that a singular matrix can be written as a product of two matrices with prescribed eigenvalues.

We now fix some notation and terminology. The set of all $n \times n$ matrices over a field F is denoted by $M_n(F)$. The determinant of a matrix A is denoted by $\det A$. The eigenvalues of a matrix A , denoted $\text{Eig } A$, are always repeated according to algebraic multiplicity, *i.e.* multiplicity as zeros of the characteristic polynomial. The nullity of A , *i.e.* the dimension of the null space is denoted by $\text{null}(A)$. The transpose of matrix A is denoted by A^t . Vectors in $F^{(n)}$ will be understood to be column vectors and are identified with $n \times 1$ matrices.

We now state our main result.

THEOREM 1. *Let A be an $n \times n$ singular matrix over a field F and let β_j and γ_j ($1 \leq j \leq n$) be elements of F . If A is not a nonzero 2×2 nilpotent, then A can be factored as a product BC with $\text{Eig } B = \{\beta_1, \dots, \beta_n\}$ and $\text{Eig } C = \{\gamma_1, \dots, \gamma_n\}$ if and only if the number of zeros m among $\beta_1, \gamma_1, \dots, \beta_n, \gamma_n$ is not less than the nullity of A . If A is a nonzero 2×2 nilpotent then A can be factored as above if and only if $1 \leq m \leq 3$.*

We note that the two extremal cases of this result are known. It was observed by C.R. Johnson (private communication) that an adaptation of the proof in [5] establishes a factorization as above with $m = \text{null } A$. This is included in the version of the first

author's factorization theorem [5] presented in the book [2]. The case $m = 2n$ is also known [3], [6], [7], (see Theorem 2 below).

PRELIMINARIES

LEMMA 1. *If A is a square matrix which is not a scalar multiple of the identity and if $\lambda \in \mathbb{F}$, then A is similar to a matrix whose (1,1) entry is λ .*

The easy proof is omitted (see also [5]).

LEMMA 2. *If $T = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ is an $n \times n$ matrix where A and D are square matrices and if A is invertible, then $\text{null}(T) = \text{null}(D - CA^{-1}B)$.*

Proof. Let $S = D - CA^{-1}B$, then

$$T = \begin{bmatrix} A & O \\ C & I \end{bmatrix} \begin{bmatrix} I & A^{-1}B \\ O & S \end{bmatrix}.$$

Since the factor on the left is invertible, it follows that

$$\text{rank}(T) = \text{rank} \begin{bmatrix} I & A^{-1}B \\ O & S \end{bmatrix} = n - \text{null}(S) \text{ and so } \text{null}(T) = \text{null}(S). \quad \blacksquare$$

LEMMA 3. *Let D be an invertible $n \times n$ matrix, and Y any $k \times n$ matrix, then the two $(n+k) \times (n+k)$ matrices*

$$\begin{bmatrix} O & O \\ O & D \end{bmatrix} \text{ and } \begin{bmatrix} O & Y \\ O & D \end{bmatrix}$$

are similar.

Proof.

$$\begin{bmatrix} \mathbf{I} & \mathbf{YD}^{-1} \\ \mathbf{O} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{I} & -\mathbf{YD}^{-1} \\ \mathbf{O} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{O} & \mathbf{Y} \\ \mathbf{O} & \mathbf{D} \end{bmatrix}. \quad \blacksquare$$

The next result was first noted in [1]. We give a proof; for another proof see [7].

LEMMA 4. *If A is a nonzero 2×2 nilpotent matrix, then A is not a product of two nilpotent matrices.*

Proof. By applying a similarity transformation, we may assume, without loss of generality, that $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$. If $A = BC$ where B and C are nilpotents, then we must have that $\text{rank } B = \text{rank } C = \text{rank } A = 1$. Furthermore $\mathcal{N}(C) \subseteq \mathcal{N}(A)$, where $\mathcal{N}(T)$ denotes the null space of T . Since both A and C have rank one, it follows that $\mathcal{N}(C) = \mathcal{N}(A) = \text{span}\{e_2\}$, where $e_2 = (0,1)^t$. Similarly $\text{range } B = \text{range } A = \text{span}\{e_2\}$. Therefore $C = \begin{bmatrix} 0 & 0 \\ c & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 \\ b & 0 \end{bmatrix}$ for some b and c in F . But then $BC = 0$, a contradiction. \blacksquare

We make use of the following result about products of nilpotent matrices. For matrices over the field of complex numbers, it was proved by Wu [7]; for arbitrary fields, two independently obtained proofs are in [3] and [6].

THEOREM 2 (The nilpotent factorization theorem). *Let A be a singular square matrix over an arbitrary field. Then A is a product of two nilpotent matrices if and only if A is not a nonzero 2×2 nilpotent matrix.*

PROOF OF THEOREM 1. We employ the following notation. We denote the list $(\beta_1, \dots, \beta_n)$ by β , the list $(\gamma_1, \dots, \gamma_n)$ by γ , and $(\beta_1, \dots, \beta_n, \gamma_1, \dots, \gamma_n)$ by $\beta \cup \gamma$. We say that $\beta = \mathbf{0}$, if $\beta_1 = \beta_2 = \dots = \beta_n = 0$. We also denote the reduced list $(\beta_2, \dots, \beta_n)$ by β' .

Necessity. Assume that $A = BC$ with $\text{Eig } B = \beta$ and $\text{Eig } C = \gamma$. Let m_1 and m_2 denote the number of zeros among $\{\beta_1, \dots, \beta_n\}$ and $\{\gamma_1, \dots, \gamma_n\}$ respectively, then

$$\text{null}(A) \leq \text{null}(B) + \text{null}(C) \leq m_1 + m_2 = m.$$

Furthermore, if A is a 2×2 nonzero nilpotent, we have $m \geq 1$ by the above and $m \neq 4$ by Lemma 4, so $1 \leq m \leq 3$.

Sufficiency. We proceed by induction on the size of the matrix A . The result is trivial for 1×1 matrices. Next assume that the conclusion of the theorem is true for all square matrices with size less than n , and let A be an $n \times n$ singular matrix and $\beta_1, \gamma_1, \dots, \beta_n, \gamma_n$ be elements of F , exactly m of which are zero where m satisfies the conditions of Theorem 1. We consider two cases according as $m < n$ or $n \leq m \leq 2n - 1$; the case $m = 2n$ being already known (Theorem 2).

Case 1. $m < n$. In this case, $\beta \neq \mathbf{0}$ and $\gamma \neq \mathbf{0}$, so with no loss of generality we assume that $\beta_1 \neq 0$, $\gamma_1 \neq 0$. The matrix A is nonzero since $\text{null}(A) \leq m < n$; thus, by Lemma 1, A is similar to the matrix

$$A_1 = \begin{bmatrix} \beta_1 \gamma_1 & y^t \\ x & D \end{bmatrix}$$

where $x, y \in F^{(n-1)}$ and $D \in M_{n-1}(F)$. It suffices to establish the desired factorization for A_1 . The number of zeros among $\beta' \cup \gamma'$ equals m and by Lemma 2,

$\text{null}(D - \beta_1^{-1} \gamma_1^{-1} x y^t) = \text{null}(A_1) < m$, we may apply the induction hypothesis to establish the existence of matrices B_o and C_o such that $D - \beta_1^{-1} \gamma_1^{-1} x y^t = B_o C_o$ and $\text{Eig } B_o = \beta'$, $\text{Eig } C_o = \gamma'$. (We notice that the exceptional 2×2 case is not present here, since if $n = 3$ and $\beta_2 = \beta_3 = \gamma_2 = \gamma_3 = 0$, then $m = 4$ and so $m > n$). Now

$$A_1 = \begin{bmatrix} \beta_1 & 0 \\ \gamma_1^{-1} x & B_o \end{bmatrix} \begin{bmatrix} \gamma_1 & \beta_1^{-1} y^t \\ 0 & C_o \end{bmatrix}$$

as required.

Case 2. $n \leq m \leq 2n - 1$. We may assume that $\beta \neq 0$, since otherwise we may factor A^t as $A^t = RS$ with $\text{Eig } R = \gamma$, $\text{Eig } S = \beta$, and then obtain $A = S^t R^t$. So we may take $\beta_1 \neq 0$. Furthermore $m \geq n$, so at least one $\gamma_j = 0$ and we therefore take $\gamma_1 = 0$.

The matrix A is similar to the matrix

$$A_1 = \begin{bmatrix} 0 & y^t \\ 0 & D \end{bmatrix}$$

where $y \in F^{(n-1)}$ and $D \in M_{n-1}(F)$. We notice that the number of zeros among $\beta' \cup \gamma'$ is $m - 1 \geq n - 1 \geq \text{null } D$. We consider two subcases.

Case 2(a). **D is nonsingular.** In this case, there exists a rank-one matrix R such that $D - R$ is singular and in case $n = 3$, $D - R$ is not nilpotent. (For example we may take R to agree with D in one row and have all other rows zero, except when D is a 2×2 matrix with zero diagonal, in which case we first apply a similarity transformation to replace D by a diagonal matrix similar to D and then proceed as before.) The number of zeros among $\beta' \cup \gamma'$ is $m - 1$ and $m - 1 \geq n - 1 \geq \text{null}(D - R)$; therefore, by the

induction hypothesis, $D - R = B_o C_o$ with $\text{Eig } B_o = \beta'$ and $\text{Eig } C_o = \gamma'$. The rank-one matrix R can be written in the form zw^t for some z and $w \in F^{(n-1)}$. Let

$$B = \begin{bmatrix} \beta_1 & 0 \\ z & B_o \end{bmatrix} \text{ and } C = \begin{bmatrix} 0 & w^t \\ 0 & C_o \end{bmatrix},$$

then $\text{Eig } B = \beta$, $\text{Eig } C = \gamma$, and

$$BC = \begin{bmatrix} 0 & \beta_1 w^t \\ 0 & D \end{bmatrix}$$

which is similar to A_2 , by Lemma 3.

Case 2(b). D is singular. Unless A is a 3×3 nonzero nilpotent and

$\beta_2 = \gamma_2 = \beta_3 = \gamma_3 = 0$, we may apply the induction hypothesis to D to establish the existence of matrices B_o and C_o such that $D = B_o C_o$ and $\text{Eig } B_o = \beta'$ and $\text{Eig } C_o = \gamma'$. The matrix A_1 can now be factored as follows:

$$A_1 = \begin{bmatrix} \beta_1 & 0 \\ 0 & B_o \end{bmatrix} \begin{bmatrix} 0 & \beta_1^{-1} y^t \\ 0 & C_o \end{bmatrix}.$$

If A is a 3×3 nilpotent of rank 1, then A is similar to $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and we proceed as before with $D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

If A is a 3×3 nilpotent of rank 2 and $\beta_2 = \beta_3 = \gamma_1 = \gamma_2 = \gamma_3 = 0$, then A is similar to a 3×3 upper triangular Jordan cell which has the following factorization:

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & \beta_1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

This ends the proof of Theorem 1. ■

PRODUCTS OF HERMITIAN AND POSITIVE SEMIDEFINITE MATRICES

We use Theorem 1 and the main result in [5] to give short proofs of and slightly generalize results of Radjavi and Wu. The result of part (a) in the next theorem is due to Wu [8]. In [4], Radjavi proved that every complex square matrix with real determinant is a product of four hermitian matrices, so parts (b) and (c) of Theorem 3 below give a generalization.

THEOREM 3. *Let A be a real or complex square matrix.*

(a) *If A is singular, then A is a product of four positive semidefinite matrices, three of which may be taken to be definite.*

(b) *If $\det A$ is a nonzero real number and A is not a scalar multiple of I , then A is a product of four hermitian matrices, at least three of which may be taken to be positive definite.*

(c) *If $A = \lambda I$, and $\det A$ is real, then A is a product of four hermitian matrices none of which can be definite unless λ^2 is real.*

REMARK. If A is a real matrix, then A is a product of two real hermitian matrices; indeed every matrix over a field is a product of two symmetric matrices over the same field. This is a classical result of Frobenius. Therefore part (b) is of interest mainly for complex matrices.

PROOF OF THEOREM 3. Each of the classes of hermitian, positive definite and positive semidefinite matrices is invariant under congruence $T \mapsto C^*TC$ for invertible C . The equation

$$T^{-1}R_1R_2R_3R_4T = (T^{-1}R_1T^{*-1})(T^*R_2T)(T^{-1}R_3T^{*-1})(T^*R_4T)$$

shows that the set of products described in the statement of Theorem 3 is invariant under similarity.

(a) First, assume that A is singular. Using standard canonical forms, it is easy to see that A is similar to a direct sum of matrices each of which has nullity one. Therefore it suffices to prove the conclusion of the theorem when $\text{null}(A) = 1$. In this case, by Theorem 1, there exist matrices B (respectively C) such that $A = BC$ and the eigenvalues of B are distinct and positive (respectively nonnegative). Therefore B (respectively C) is similar to a positive definite matrix (respectively positive semidefinite), so $B = T^{-1}PT$, $C = R^{-1}SR$ with P positive definite, S positive semidefinite, and T and R invertible. The equation

$$A = (T^{-1}PT^{*-1})(T^*T)(R^{-1}R^{*-1})(R^*SR)$$

gives A as a product of three positive definite matrices and one positive semidefinite.

(b) If $\det T$ is a nonzero real, and T is not a scalar multiple of I , we use the main theorem in [5] to write A as a product of matrices BC where B has n distinct positive eigenvalues and C has n distinct real eigenvalues at most one of which is negative. As before B is a product of two positive definite matrices and C is a product of a positive definite and a hermitian matrix H . Therefore A is a product of four hermitian matrices, at least three of which can be taken positive definite. We note that if $\det A > 0$, the above proof gives A as a product of four positive definite matrices, a result of Ballantine (see [5]).

(c) If $A = \lambda I$, the proof is as given in [4]. Choose a basis $\{e_1, \dots, e_n\}$ of $F^{(n)}$. Let U be the shift given by $Ue_j = e_{j+1}$ for $1 \leq j \leq n-1$ and $Ue_n = e_1$. Therefore $A = U \cdot \lambda U^t$. If S is the hermitian symmetry given by $Se_j = e_{n+1-j}$ and K the operator given by $Ke_j = \bar{\lambda}^{n+1-j} \lambda^j e_{n+1-j}$, then $US, S, \lambda U^t K^{-1}$ and K are hermitian and $A = (US)S(\lambda U^t K^{-1})K$, a product of four hermitian matrices.

Furthermore if λI is a product of four hermitian matrices at least one of which is positive definite, then $\lambda PH = H_1 H_2$ for hermitian H_1, H_2, H and a positive definite P . Now $H_1 H_2$ is similar to $H_2 H_1 = (H_1 H_2)^*$, so λPH is similar to its adjoint $\bar{\lambda} HP$, which in turn is similar to $\bar{\lambda} PH$. Therefore $\lambda^2 PH$ is similar to $\lambda \bar{\lambda} PH$. But PH has real eigenvalues, since it is similar to the positive definite matrix $P^{\frac{1}{2}} H P^{\frac{1}{2}}$. Therefore λ^2 is real. ■

REMARK 1. If λ^2 is real, then λI is either hermitian or is a product of three hermitian matrices. This follows from the equation

$$\begin{bmatrix} i & 0 \\ 0 & i \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

REMARK 2. The position of the three positive definite matrices in parts (a) and (b) of Theorem 3 is quite arbitrary. This follows from the fact that a product PH of a positive definite P and hermitian (respectively positive semidefinite) H can be written as a product of two matrices of the same type but in reverse order as $PH = (PHP^*)P^{-1}$.

REFERENCES

- [1] C.K. Fong and A.R. Sourour, Sums and products of quasi-nilpotent operators, *Proc. Roy. Soc. Edinburgh* **99A**(1984), 193–200.
- [2] R. Horn and C. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
- [3] T.J. Laffey, Products of matrices. *Preprint*.
- [4] H. Radjavi, Products of hermitian matrices and symmetries, *Proc. Amer. Math. Soc.* **21**(1969), 369–372; Errata, *Proc. Amer. Math. Soc.* **26**(1970), 701.
- [5] A.R. Sourour, A factorization theorem for matrices, *Linear and Multilinear Algebra* **19**(1986), 141–147.
- [6] A.R. Sourour, Nilpotent factorization of matrices. *Preprint*.
- [7] P.Y. Wu, Products of nilpotent matrices, *Linear Algebra Appl.* **96**(1987), 227–232.
- [8] P.Y. Wu, Products of positive semidefinite matrices, *Linear Algebra Appl.* **111**(1988), 53–61.
- [9] P.Y. Wu, The operator factorization problems, *Linear Algebra Appl.* **117**(1989), 35–63.

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF VICTORIA, VICTORIA, B.C.
V8W 3P4, CANADA.